

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: November 15, 2005, 23:53:58 / Search time 31202 Seconds  
(without alignments)

11774.484 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7582

Sequence: 1 caacaatcgagataataacc.....tataaactctgcactgcr 7582

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenBank:  
1: gb\_ba:  
2: gb\_hcg:  
3: gb\_in:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_pro:  
11: gb\_sce:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vi:

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7489.6	98.8	7582	BD196245	BD196245 Endogenet
2	7489.6	98.8	7582	BD267487	BD267487 Endogenet
3	7489.6	98.8	7582	AX000966	AX000966 Sequence
4	7489.6	98.8	7582	AX027480	AX027480 Sequence
5	6448	85.0	10222	AY101582	AY101582 Homo sapi
6	6448	85.0	10222	AY101585	AY101585 Homo sapi
7	6448	85.0	56093	AX329572	AX329572 Sequence
8	6448	85.0	56093	HSAC000064	HSAC000064 Human BAC
9	6448	85.0	149194	AC007565	AC007565 Homo sapi
10	6446.4	85.0	10222	AY101583	AY101583 Homo sapi
11	6440	84.9	10222	AY101584	AY101584 Homo sapi
12	6436	84.9	10499	BD221808	BD221808 Nucleic s
13	6436	84.9	10499	AX007980	AX007980 Sequence
14	6346.4	83.7	10229	AY101586	AY101586 Pan trogl
15	6344.8	83.7	10229	AY101587	AY101587 Pan trogl
16	6344.8	83.7	184675	AC145964	AC145964 Pan trogl
17	6269.6	82.7	10230	AY101588	AY101588 Gorilla g
18	6269.6	82.7	10230	AY101589	AY101589 Gorilla g
19	6037.8	79.6	10122	AY101590	AY101590 Pongo pyg

20	6025.8	79.5	10124	9	AY101591	AY101591 Pongo pyg
21	5956.8	78.6	10246	9	AY101593	AY101593 Hylobates
22	5950.4	78.5	10248	9	AY101592	AY101592 Hylobates
23	5824.2	76.8	158033	9	AY1018926	AY1018926 Homo sapi
24	5492	72.4	105989	9	AC008121	AC008121 Homo sapi
25	5492	72.4	110000	2	AC009727_2	AC009727_2 Homo sapi
26	4755.6	62.7	187321	2	AC092510	AC092510 Papio anu
27	4302.2	56.7	114621	9	AC005187	AC005187 Homo sapi
28	4295.8	56.7	46575	2	AC080036	AC080036 Homo sapi
29	4162	54.9	111140	9	AL592310	AL592310 Homo sapi
30	3944.6	52.0	183680	9	AC098859	AC098859 Homo sapi
31	3779.8	49.9	73070	9	AL590143	AL590143 Homo sapi
32	3731.6	49.2	173788	9	AC022555	AC022555 Homo sapi
33	3731.6	49.2	186540	9	AC090341	AC090341 Homo sapi
34	3731.6	49.2	187837	9	AC079065	AC079065 Homo sapi
35	3716.6	49.0	160883	2	AL158814	AL158814 Homo sapi
36	3537.2	46.7	137947	9	HS4514	HS4514 Homo sapi
37	3442.8	45.4	153444	2	AC040948	AC040948 Homo sapi
38	3442.8	45.4	205035	9	CNS00005	AL049870 Homo chr
39	3438.2	45.3	101846	9	HS19711	AL031390 Human DNA
40	3438.2	45.3	224629	2	AC139260	AC139260 Homo sapi
41	3420.6	45.1	68693	9	AC108746	AC108746 Homo sapi
42	3367.4	44.4	118312	9	AC055713	AC055713 Homo sapi
43	3367.4	44.4	197549	2	AC145423	AC145423 Homo sapi
44	3300.6	43.5	169250	9	AP002812	AP002812 Homo sapi
45	3290.2	43.4	134412	2	AP002377	AP002377 Homo sapi

## ALIGNMENTS

RESULT 1	BD196245	7582 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD196245				
DEFINITION	Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders.				
ACCESSION	BD196245.1	GI:33006015			
VERSION	BD196245.1	GI:33006015			
KEYWORDS	JP 2002512530-A/11.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 7582)				
AUTHORS	Beseme, F., Blond, J.L., Bouton, O., Mandrand, B. and Mallet, F.				
TITLE	Endogenetic retroviral sequences, associated with autoimmune diseases and/or with pregnancy disorders				
JOURNAL	Patent: JP 2002512530-A 11 23-APR-2002;				
COMMENT	BIO MERIEUX				
OS	Unidentified				
PN	JP 2002512530-A/11				
PD	23-APR-2002				
PR	06-SEP-1998 JP 1999508244				
PR	07-JUL-1997 FR 97/08815				
PI	FREDERIC BESEME, JEAN LUC BLOND, OLIVIER BOUTON, BERNARD MANDRAND,				
PI	FRANCOIS MALLET				
PC	C12N15/48, C07K14/15, C12Q1/68, C07K16/10, G01N33/569 CC				
Strandness:	Single;				
CC	Topology: Linear;				
CC	Endogenetic retroviral sequences, associated with autoimmune diseases				
CC	diseases				
CC	and/or with pregnancy disorders				
FH	Key				
FT	source				
FT	1. 7582				
FEATURES	Location/Qualifiers				
source	1. 7582				
	/organism="Unidentified".				
	/mol_type="unidentified"				
	/db_xref="taxon:32644"				
ORIGIN					
Query Match	98.8%; Score 7489.6; DB 6; Length 7582;				

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 CAACAATCGGGAATATAACCCAGGCAATGAGCTGCGAACAAGACCCCTTTGGGTCC 60
DB      1 CAACAATCGGGAATATAACCCAGGCAATGAGCTGCGAACAAGACCCCTTTGGGTCC 60
QY      61 CTTCCCTTTGTATAGGGAGCTGTTTCAATGATTTCACTATTTCACTATTAATCTTGAACGCA 120
DB      61 CTTCCCTTTGTATAGGGAGCTGTTTCAATGATTTCACTATTTCACTATTAATCTTGAACGCA 120
QY      121 CTTCTTGATGTCATGTTTCTTAAGGCTGAGCTGAGCTGTTTGTCTCAACCTGCACTGTC 180
DB      121 CTTCTTGATGTCATGTTTCTTAAGGCTGAGCTGAGCTGTTTGTCTCAACCTGCACTGTC 180
QY      181 TGTGTTCCACACCGCAGCAGCAGCTGCGCTGACCTCCCATTTGGGCTAAAGGCTGTC 240
DB      181 TGTGTTCCACACCGCAGCAGCAGCTGCGCTGACCTCCCATTTGGGCTAAAGGCTGTC 240
QY      241 CGGTGCTCCGATTCAGAGGAGCGGCTGATGCTCCCATTTGGGCTAAAGGCTGTC 300
DB      241 CGGTGCTCCGATTCAGAGGAGCGGCTGATGCTCCCATTTGGGCTAAAGGCTGTC 300
QY      301 CATGTTNCTGACAGGCTAGTGCCTGGTGTGTTCTAATTGAGCTGACACTAATCACT 360
DB      301 CATGTTNCTGACAGGCTAGTGCCTGGTGTGTTCTAATTGAGCTGACACTAATCACT 360
QY      361 GGGTTCATGTTCTCTTCTGTGACCGAGGCTTCTAATAKAACTATPAACATTTACCACA 420
DB      361 GGGTTCATGTTCTCTTCTGTGACCGAGGCTTCTAATAKAACTATPAACATTTACCACA 420
QY      421 TGGCCCAAGTTCATTTCTTGGAAATCCGTAAGGSSACAAGAACTCCAGGCTCAAGAAATAC 480
DB      421 TGGCCCAAGTTCATTTCTTGGAAATCCGTAAGGSSACAAGAACTCCAGGCTCAAGAAATAC 480
QY      481 GAGCTTGGCCACCATCTTGAAGGCGCTGCTACCTCTTGAAGTGTTCACCAACCATC 540
DB      481 GAGCTTGGCCACCATCTTGAAGGCGCTGCTACCTCTTGAAGTGTTCACCAACCATC 540
QY      541 TTGGAGCTCTGTGAGCAAGAGACCCCGGTACATTTTGGCAACAMSPAAGGACATCC 600
DB      541 TTGGAGCTCTGTGAGCAAGAGACCCCGGTACATTTTGGCAACAMSPAAGGACATCC 600
QY      601 MAAGTATGAGGAAACGTTCCCGCAACAACAAAACGCGCTTAAGACCTAATCTTGGAAAT 660
DB      601 MAAGTATGAGGAAACGTTCCCGCAACAACAAAACGCGCTTAAGACCTAATCTTGGAAAT 660
QY      661 TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAAGAACTTATATTCTTCTGAGTGC 720
DB      661 TGGGAMCAATTTGACCTCAGACACTAAGAAAGAAAGAACTTATATTCTTCTGAGTGC 720
QY      721 GCTGAGCACTCTGAGGAAAGTATTAATTATACACATCTTACAGCTAAGCYTCTTTTG 780
DB      721 GCTGAGCACTCTGAGGAAAGTATTAATTATACACATCTTACAGCTAAGCYTCTTTTG 780
QY      781 TAGAAAAGGCAATGAGTGAAGTCCATTAAGTCAAACTTTCTTTCACTTAAGACAA 840
DB      781 TAGAAAAGGCAATGAGTGAAGTCCATTAAGTCAAACTTTCTTTCACTTAAGACAA 840
QY      841 CTACAAATTAATGTAAGAGTGAATTAAGCCTACAGAAAGCTTCAAGATCTACTCC 900
DB      841 CTACAAATTAATGTAAGAGTGAATTAAGCCTACAGAAAGCTTCAAGATCTACTCC 900
QY      901 CTAATCCAGCATCCCGCATCTCCTTCCCAATTAAGAGACCCCTTCAACCCAAATGG 960
DB      901 CTAATCCAGCATCCCGCATCTCCTTCCCAATTAAGAGACCCCTTCAACCCAAATGG 960
QY      961 TCCAAAAGAGATGACAAAAGGTTAAACATGAACCAAAGATGCCAATTTCCCAAT 1020
DB      961 TCCAAAAGAGATGACAAAAGGTTAAACATGAACCAAAGATGCCAATTTCCCAAT 1020
QY      1021 TATGACCCCTCCCAAGAGTGGGAGGAAGAAATTCGGCCAGGCAAGTGCATGTGCT 1080
DB      1021 TATGACCCCTCCCAAGAGTGGGAGGAAGAAATTCGGCCAGGCAAGTGCATGTGCT 1080

```

```

DB      1081 TTTTCTCCAGACTTAAGCAAAATTAAGAACTTAAGTAAATTTCCAGATAATCTT 1140
QY      1081 TTTTCTCCAGACTTAAGCAAAATTAAGAACTTAAGTAAATTTCCAGATAATCTT 1140
DB      1141 GATGCTAATTAATGTTTATTAAGGGTTAGAACAACTTTGATCTGACATGAGAGAT 1200
QY      1141 GATGCTAATTAATGTTTATTAAGGGTTAGAACAACTTTGATCTGACATGAGAGAT 1200
DB      1201 ATATATGTAATGCTTAATTAACACACTAACCCCAATGAGAGGCAACCTAATCTGC 1260
QY      1201 ATATATGTAATGCTTAATTAACACACTAACCCCAATGAGAGGCAACCTAATCTGC 1260
DB      1261 AGCTGAGAGTTGGGATCTCTGATCTCAGTCAAGTCAATGATTAAGATGACACA 1320
QY      1261 AGCTGAGAGTTGGGATCTCTGATCTCAGTCAAGTCAATGATTAAGATGACACA 1320
DB      1321 GAAGGAAGANATGATTTCCCAAGGCGACAGACGAGTTCCAGTCTAASCCCTCAT 1380
QY      1321 GAAGGAAGANATGATTTCCCAAGGCGACAGACGAGTTCCAGTCTAASCCCTCAT 1380
DB      1381 GGGACACAGAAATCAGTAAATGAGAGATTTGGTCTGACAGATTGGCTTAATCTGTGTC 1440
QY      1381 GGGACACAGAAATCAGTAAATGAGAGATTTGGTCTGACAGATTGGCTTAATCTGTGTC 1440
DB      1441 TASAAGGCTAAGAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCT 1500
QY      1441 TASAAGGCTAAGAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCT 1500
DB      1501 AACAGAGGAGAGAGAGAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCT 1560
QY      1501 AACAGAGGAGAGAGAGAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCTAAGAAATCT 1560
DB      1561 GAAAGCTGCTCTCTGACCTGACCTCTTGAAGGCACTAATCTTAAGGTAAGT 1620
QY      1561 GAAAGCTGCTCTCTGACCTGACCTCTTGAAGGCACTAATCTTAAGGTAAGT 1620
DB      1621 TATCACTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1680
QY      1621 TATCACTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1680
DB      1681 CAAAACCTAAGAAACCTAATGAACCTTGGCAACTGCGTTTATTAATAGATCAGAG 1740
QY      1681 CAAAACCTAAGAAACCTAATGAACCTTGGCAACTGCGTTTATTAATAGATCAGAG 1740
DB      1741 GAGCAGCGGACAGACAAACGAGATTAAGAAAGGCAACGCTTATCTATGACCT 1800
QY      1741 GAGCAGCGGACAGACAAACGAGATTAAGAAAGGCAACGCTTATCTATGACCT 1800
DB      1801 CAGGCAAGGAGACTTTGAGAGCTCTGAAAAGGAAAGCTGGGCAATTTGAATGCTTA 1860
QY      1801 CAGGCAAGGAGACTTTGAGAGCTCTGAAAAGGAAAGCTGGGCAATTTGAATGCTTA 1860
DB      1861 TAGGCTGCTTCAAGTGGGCTTAAGAGCACTTTAAAGATTTGCTCAATGAAG 1920
QY      1861 TAGGCTGCTTCAAGTGGGCTTAAGAGCACTTTAAAGATTTGCTCAATGAAG 1920
DB      1921 TAAGCGGCTCTGCTGATGCTGCTTATTAAGGAAACACTGGAAGGCGCACTGCCC 1980
QY      1921 TAAGCGGCTCTGCTGATGCTGCTTATTAAGGAAACACTGGAAGGCGCACTGCCC 1980
DB      1981 CAGGAGGCAAAAGGCTTTTGAATCAGAGCACTAACCAGATATCCAGAGCAGAGCTG 2040
QY      1981 CAGGAGGCAAAAGGCTTTTGAATCAGAGCACTAACCAGATATCCAGAGCAGAGCTG 2040
DB      2041 AGGATGCTGGGAGAGGCAATCCCATGCAATCCCTCAAGAGCCTGGGATGCTT 2100
QY      2041 AGGATGCTGGGAGAGGCAATCCCATGCAATCCCTCAAGAGCCTGGGATGCTT 2100
DB      2101 GACCAATTGAGGCGCAGAGAGGTTGCTCTGAGACCTGAGGCTTCTTAATGCTTAAT 2160
QY      2101 GACCAATTGAGGCGCAGAGAGGTTGCTCTGAGACCTGAGGCTTCTTAATGCTTAAT 2160
DB      2101 GACCAATTGAGGCGCAGAGAGGTTGCTCTGAGACCTGAGGCTTCTTAATGCTTAAT 2160

```

OY 2161 CTTCTGTCGGGACAACTGTCTCCAGATCTGTCACTATTCTGAGGGGTCCTTAAGACG 2220  
Db 2161 CTTCTGTCCCGGACAACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGTCCTTAAGACG 2220  
OY 2221 GGCAGTCACTAGATATCTTTTCCCGCACAATTAGTAACTGGGGAGCTTTATCTTT 2280  
Db 2221 GGCAGTCACTAGATATCTTTTCCCGCACAATTAGTAACTGGGGAGCTTTATCTTT 2280  
OY 2281 TCACATGCTTTTCTAATTATGCTTGAAGCCCACTACCTGTTTGAAGAGACATTTCTA 2340  
Db 2281 TCACATGCTTTTCTAATTATGCTTGAAGCCCACTACCTGTTTGAAGAGACATTTCTA 2340  
OY 2341 GCAAAAGCAGGGGCCATTATACCTCTGAACATAGAGAGAGAACACCCGTTGTTGTC 2400  
Db 2341 GCAAAAGCAGGGGCCATTATACCTCTGAACATAGAGAGAGAACACCCGTTGTTGTC 2400  
OY 2401 CCTGCTTGAGGAAGGAATTAATCTGTAAGTCTGGGCAACAGAGACAAATATGACGAGC 2460  
Db 2401 CCTGCTTGAGGAAGGAATTAATCTGTAAGTCTGGGCAACAGAGACAAATATGACGAGC 2460  
OY 2461 CAAAGATGCCCCCTGCTGTCTGAATTAACCTAAGAGATCCACTTCCCTTCCCAACAA 2520  
Db 2461 CAAAGATGCCCCCTGCTGTCTGAATTAACCTAAGAGATCCACTTCCCTTCCCAACAA 2520  
OY 2521 GGCAGTACCCCTCTCAGACCCCAAGGCCCAAGAGATCCAAAGATTTGAAGACTTAA 2580  
Db 2521 GGCAGTACCCCTCTCAGACCCCAAGGCCCAAGAGATTTGAAGACTTAA 2580  
OY 2581 AAGCCCAAGGCTTATGTAACCATGCACTAATCTCTGCAAGTAACTCCGTAGTGAGTTGAG 2640  
Db 2581 AAGCCCAAGGCTTATGTAACCATGCACTAATCTCTGCAAGTAACTCCGTAGTGAGTTGAG 2640  
OY 2641 GAGGCAAGAAACCCAGTGAAGAGAGGTTAGTCAAGATTCAGAGATTAATCAATG 2700  
Db 2641 GAGGCAAGAAACCCAGTGAAGAGAGGTTAGTCAAGATTCAGAGATTAATCAATG 2700  
OY 2701 AAGCCGTTGTCTTTTATACCCAGCTGTACTAGCCCTTATACCTGTGMYTTCCTCAATAC 2760  
Db 2701 AAGCCGTTGTCTTTTATACCCAGCTGTACTAGCCCTTATACCTGTGMYTTCCTCAATAC 2760  
OY 2761 CAGAGAAACAGAGTGTGTTTACASTCTCTGACCTTMAAGATGCTTCTTGTGATCCCTG 2820  
Db 2761 CAGAGAAACAGAGTGTGTTTACASTCTCTGACCTTMAAGATGCTTCTTGTGATCCCTG 2820  
OY 2821 TACATCTGACTCTCAATCTTGTGCTTGAAGTACTTGAACCCCACTCAAC 2880  
Db 2821 TACATCTGACTCTCAATCTTGTGCTTGAAGTACTTGAACCCCACTCAAC 2880  
OY 2881 TCACCTGCACTRTTTTATACCCCAAGGTTCAAGGATAGCCCCATCTAATTGGCCAGGAT 2940  
Db 2881 TCACCTGCACTRTTTTATACCCCAAGGTTCAAGGATAGCCCCATCTAATTGGCCAGGAT 2940  
OY 2941 TAGCCCAAGACTTGAGCACTTGTGCTTGAAGTACTTGAACCCCACTCAAC 3000  
Db 2941 TAGCCCAAGACTTGAGCACTTGTGCTTGAAGTACTTGAACCCCACTCAAC 3000  
OY 3001 ATTACTTTTTRGCGCTTTCAGAACTTGTGCTTGAAGTACTTGAACCCCACTCAAC 3060  
Db 3001 ATTACTTTTTRGCGCTTTCAGAACTTGTGCTTGAAGTACTTGAACCCCACTCAAC 3060  
OY 3061 ATTCTCTGCGTACTGTGCTTCAWAGTTTCCAAASABARGTCTCTGTCTCAGAC 3120  
Db 3061 ATTCTCTGCGTACTGTGCTTCAWAGTTTCCAAASABARGTCTCTGTCTCAGAC 3120  
OY 3121 AGGTAAATTAATTAAGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180  
Db 3121 AGGTAAATTAATTAAGTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180  
OY 3181 TCAGAGCTAATCTGCTTATCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3240  
Db 3181 TCAGAGCTAATCTGCTTATCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACT 3240

OY 3241 GCRTAAYAGGTTTCTGCCGAATATGATTTCCCAAGTGTGGCBAATATGCCAGCTCATTA 3300  
Db 3241 GCRTAAYAGGTTTCTGCCGAATATGATTTCCCAAGTGTGGCBAATATGCCAGCTCATTA 3300  
OY 3301 WATACASTAATTAAGAACTCAGAAAGCCATATCCATTAATTAATTAATTAATTAATTAAT 3360  
Db 3301 WATACASTAATTAAGAACTCAGAAAGCCATATCCATTAATTAATTAATTAATTAATTAAT 3360  
OY 3361 GYRBAATGCTTCCAGGCCCCCTTAAAGAGCCCTTAAACCAAGCTTCACTGTTAAGT 3420  
Db 3361 GYRBAATGCTTCCAGGCCCCCTTAAAGAGCCCTTAAACCAAGCTTCACTGTTAAGT 3420  
OY 3421 TGCCACRGGGCAAGATTTTSTTATATATCAAGAAAAAAGABAAVAGCTTREGA 3480  
Db 3421 TGCCACRGGGCAAGATTTTSTTATATATCAAGAAAAAAGABAAVAGCTTREGA 3480  
OY 3481 GTCCTTACACAGTCCBAGGAGAGAGCTTGCACCTATGCGRYACTGASTAAGAAAT 3540  
Db 3481 GTCCTTACACAGTCCBAGGAGAGAGCTTGCACCTATGCGRYACTGASTAAGAAAT 3540  
OY 3541 GATGTAGTGGCAAGGTTGCTTCACTTGTTTAYGGTATGTTGCGCATGACGTAT 3600  
Db 3541 GATGTAGTGGCAAGGTTGCTTCACTTGTTTAYGGTATGTTGCGCATGACGTAT 3600  
OY 3601 GTATCTGAAGCAGTTAAATTAATACAGGGAGAGATCTTATCTGTGGAACATCTCATGAK 3660  
Db 3601 GTATCTGAAGCAGTTAAATTAATACAGGGAGAGATCTTATCTGTGGAACATCTCATGAK 3660  
OY 3661 GTGAAYRGATACTCATCTGTAAAGAGACTTGTGCTGTCAAGCACTGTTACTTAA 3720  
Db 3661 GTGAAYRGATACTCATCTGTAAAGAGACTTGTGCTGTCAAGCACTGTTACTTAA 3720  
OY 3721 TTTCAAGCTCTATTAATTTAGAGGCGAGTGTGCTGCACTGTGCACTTCTTAA 3780  
Db 3721 TTTCAAGCTCTATTAATTTAGAGGCGAGTGTGCTGCACTGTGCACTTCTTAA 3780  
OY 3781 CCAGYCNCAATTTCTTCAGACATGAAAGAAAGATTAATTAATTAATTAATTAATTAAT 3840  
Db 3781 CCAGYCNCAATTTCTTCAGACATGAAAGAAAGATTAATTAATTAATTAATTAATTAAT 3840  
OY 3841 TCTCAAACTTAATGCCACTGAGGGGACTTGTAGAGTTCCYTTGACTGATCCYACCTT 3900  
Db 3841 TCTCAAACTTAATGCCACTGAGGGGACTTGTAGAGTTCCYTTGACTGATCCYACCTT 3900  
OY 3901 CAACCTGTATCTGATGGAAGTTCCTTGTGAAGAAAGAAAGCTTGAAGGAGGATGTC 3960  
Db 3901 CAACCTGTATCTGATGGAAGTTCCTTGTGAAGAAAGAAAGCTTGAAGGAGGATGTC 3960  
OY 3961 AGTGTCAGTGATTAATGGAATYTTGAAGTATCCCTCACTCCAGGAATAGTCTTA 4020  
Db 3961 AGTGTCAGTGATTAATGGAATYTTGAAGTATCCCTCACTCCAGGAATAGTCTTA 4020  
OY 4021 GCTRGCAAGAACTTAATAGCCCTTCACTTGAAGGCACTTGAAGGCACTTGAAGG 4080  
Db 4021 GCTRGCAAGAACTTAATAGCCCTTCACTTGAAGGCACTTGAAGGCACTTGAAGG 4080  
OY 4081 TATATATCAAGCTTATATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 4140  
Db 4081 TATATATCAAGCTTATATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 4140  
OY 4141 AGAAAGGAATTCCTAATTCCTGAGAGAACTATCAACATCGAGGAAGCCATTAAGAR 4200  
Db 4141 AGAAAGGAATTCCTAATTCCTGAGAGAACTATCAACATCGAGGAAGCCATTAAGAR 4200  
OY 4201 ATTATTAATGCGWGTACAGAAACCTTAAGAGGTGMACTTTACCTGCGGCTCATCA 4260  
Db 4201 ATTATTAATGCGWGTACAGAAACCTTAAGAGGTGMACTTTACCTGCGGCTCATCA 4260  
OY 4261 NAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 4320  
Db 4261 NAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 4320  
OY 4321 TGCAAGGACGAGACCTCCTCATTAAGAAATGCTTATTAATCTTCCCTTATAGAGTAATCC 4380

Db	4321	TGCAAGGACGAGCCTCCATTTGGAATGGTTATTTAACTTCCTCTAGTATGGGTAAACC	4380
Oy	4381	CTTCCGGGAAACCAAGCCCACTACTCAAGCAGAGAAAACAGATGGGAACTCCACGAG	4440
Db	4381	CTTCCGGGAAACCAAGCCCACTACTCAAGCAGAGAAAACGATGGGAACTCCACGAG	4440
Oy	4441	CAGTTTCTCCCCCTGGGACGGTTATGCACTGAAAGAGGAAATATCTTTGGCTGCAAC	4500
Db	4441	CAGTTTCTCCCCCTGGGACGGTTATGCACTGAAAGAGGAAATATCTTTGGCTGCAAC	4500
Oy	4501	TATCCAAATGAAATTAATTAAACCCTTCATCAACCTTTCACTTAGCATCATGATAGAC	4560
Db	4501	TATCCAAATGAAATTAATTAAACCCTTCATCAACCTTTCACTTAGCATCATGATAGAC	4560
Oy	4551	CCATCABATGGCCAAATCATTTATTCAGACACAGCCCTTTCAAAATATCAAGCAAT	4620
Db	4551	CCATCABATGGCCAAATCATTTATTCAGACACAGCCCTTTCAAAATATCAAGCAAT	4620
Oy	4621	AKTCAGGACCTGTGAATGTGTCABABAAAATAATCCCTGCTATTCGCCAAGCTCTTC	4680
Db	4621	AKTCAGGACCTGTGAATGTGTCABABAAAATAATCCCTGCTATTCGCCAAGCTCTTC	4680
Oy	4681	AGABAAACAAABAAACAGGCCAATTACCTGTABABABATGGGAACTGATTTTACCACAAG	4740
Db	4681	AGABAAACAAABAAACAGGCCAATTACCTGTABABABATGGGAACTGATTTTACCACAAG	4740
Oy	4741	CCCAAACTCAGGGAATTCAGTATCTACTACTGTGGGTABATATCTTACCGGTGTGGCA	4800
Db	4741	CCCAAACTCAGGGAATTCAGTATCTACTACTGTGGGTABATATCTTACCGGTGTGGCA	4800
Oy	4801	RAGGCTTCCCCCTGTAGGACAGAAAAGGCCCAAGAGTAATPAAAGGCACTAGTTCATGAA	4860
Db	4801	RAGGCTTCCCCCTGTAGGACAGAAAAGGCCCAAGAGTAATPAAAGGCACTAGTTCATGAA	4860
Oy	4861	ATAATTTCCAGATTGGCACTTCCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAG	4920
Db	4861	ATAATTTCCAGATTGGCACTTCCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAG	4920
Oy	4921	GCCACAGTAAACCCAGGAGATATCCCAAGCGTTAAGGTATACGATATCATTTCACTGCGCC	4980
Db	4921	GCCACAGTAAACCCAGGAGATATCCCAAGCGTTAAGGTATACGATATCATTTCACTGCGCC	4980
Oy	4981	TGAAGGCCACAGTCTCTCAGGGAAGGTGCAGAAAATGAATGAAAATCTCAAGACATCTA	5040
Db	4981	TGAAGGCCACAGTCTCTCAGGGAAGGTGCAGAAAATGAATGAAAATCTCAAGACATCTA	5040
Oy	5041	AAAAAGCAAAACCCAGAGAAACCCACTCAATGCGCTGTCTGTGCTTATAGCTTAA	5100
Db	5041	AAAAAGCAAAACCCAGAGAAACCCACTCAATGCGCTGTCTGTGCTTATAGCTTAA	5100
Oy	5101	AGATCTGCAACTTTTCCCCCAAAAGACAGACTTAGCCATACGAAATGCTGTATGGAAG	5160
Db	5101	AGATCTGCAACTTTTCCCCCAAAAGACAGACTTAGCCATACGAAATGCTGTATGGAAG	5160
Oy	5161	CCCTTCAATACCAATGACCTTGTGCTTACCCCAAGACAGCCCACTTAGTTCAGACATCA	5220
Db	5161	CCCTTCAATACCAATGACCTTGTGCTTACCCCAAGACAGCCCACTTAGTTCAGACATCA	5220
Oy	5221	CCTCCTTAGCCAAATATCACAAGTTCTTTAAACCTTACAAAGAACTATCCTGAGAG	5280
Db	5221	CCTCCTTAGCCAAATATCACAAGTTCTTTAAACCTTACAAAGAACTATCCTGAGAG	5280
Oy	5281	AGGGAAGAACTATTCCACCMMGTGACATGATTTATGTAAGTCCCTTCTCTAAT	5340
Db	5281	AGGGAAGAACTATTCCACCMMGTGACATGATTTATGTAAGTCCCTTCTCTAAT	5340
Oy	5341	CCCCATCCCTAGATCATCTGTGGAGAACCTTACCCAGTATTTATATACCCCACTG	5400
Db	5341	CCCCATCCCTAGATCATCTGTGGAGAACCTTACCCAGTATTTATATACCCCACTG	5400
Oy	5401	CGGTTAAAGTGCTGAGTGAAGTCTTGGATACATCACTTGAATCAAAATCTCGATAC	5460
Db	5401	CGGTTAAAGTGCTGAGTGAAGTCTTGGATACATCACTTGAATCAAAATCTCGATAC	5460

Db	5401	CGTTTAAAGTGCGTAGTGAGTCTTGGAATACATCACTTGAGTCATAAATCCTGGATAC	5460
Qy	5461	TGCCAAAGGAACCTGAAGAAATCAGAGACAACGCTAGACTATTCCTGTGAACCTCTAGAGG	5520
Db	5461	TGCCAAAGGAACCTGAAGAAATCAGAGACAACGCTAGACTATTCCTGTGAACCTCTAGAGG	5520
Qy	5521	ATTTCGCGCTGCTCTTCAACACACACACAGAGGAAAGTAACTAAATCATTAATTCCTCC	5580
Db	5521	ATTTCGCGCTGCTCTTCAACACACACACAGAGGAAAGTAACTAAATCATTAATTCCTCC	5580
Qy	5581	ATGGSCTCCCTTATCATATATTTTCTCTKTAASGTSTTTTAAOCCSTSTTTCACCTCACT	5640
Db	5581	ATGGSCTCCCTTATCATATATTTTCTCTKTAASGTSTTTTAAOCCSTSTTTCACCTCACT	5640
Qy	5641	GCACCCCCCTCATGSCGCTGATATGACACAGTAGCTCCCTCTAACCAAGAGTTTCTATGGAGA	5700
Db	5641	GCACCCCCCTCATGSCGCTGATATGACACAGTAGCTCCCTCTAACCAAGAGTTTCTATGGAGA	5700
Qy	5701	ATGCAGCGTCCCGGAAATATTGATGCCCATTCGTATAGAGTCTTSTTAAAGGAACCCC	5760
Db	5701	ATGCAGCGTCCCGGAAATATTGATGCCCATTCGTATAGAGTCTTSTTAAAGGAACCCC	5760
Qy	5761	ACCTTCACTGCCCCACCCCATATAGCCCCCGGACCTGCTATCACTCTGCACTCTTTGGATG	5820
Db	5761	ACCTTCACTGCCCCACCCCATATAGCCCCCGGACCTGCTATCACTCTGCACTCTTTGGATG	5820
Qy	5821	CATGCAAAATACCTATTAATTTGGAACAGAGAAATGATTATCTAGTTCTCCGTGAGAGATT	5880
Db	5821	CATGCAAAATACCTATTAATTTGGAACAGAGAAATGATTATCTAGTTCTCCGTGAGAGATT	5880
Qy	5881	GGAAGTCACTGTCTGTGGACTTACTTACCCAACTGGTATGTCTGATGSGGGGTGAGTT	5940
Db	5881	GGAAGTCACTGTCTGTGGACTTACTTACCCAACTGGTATGTCTGATGSGGGGTGAGTT	5940
Qy	5941	CAAGTACAGGCAAGAGAAAAACATGTAAAGAAATATCTCCCAACTCACCGGGGTACAT	6000
Db	5941	CAAGTACAGGCAAGAGAAAAACATGTAAAGAAATATCTCCCAACTCACCGGGGTACAT	6000
Qy	6001	GGCACCTCTAGACCCCTCAAGAGGACTGATCTCTCAAAACCTACATGAAACCTCTCGTACC	6060
Db	6001	GGCACCTCTAGACCCCTCAAGAGGACTGATCTCTCAAAACCTACATGAAACCTCTCGTACC	6060
Qy	6061	CATACTCGCCTGGTAGCCTATTTAATACCACTCTAGGGCTCCATGAGGTCTCGGCC	6120
Db	6061	CATACTCGCCTGGTAGCCTATTTAATACCACTCTAGGGCTCCATGAGGTCTCGGCC	6120
Qy	6121	CAAAACCTCTACTAATCTGTGGATATGCTCTCCCTGAACTTTCARGCCATATGTTTCAATC	6180
Db	6121	CAAAACCTCTACTAATCTGTGGATATGCTCTCCCTGAACTTTCARGCCATATGTTTCAATC	6180
Qy	6181	CTGTACTCCTGAACAATGGAAACAACCTTAGAGACAGAAATAAACAACACTTCGCTTATGTA	6240
Db	6181	CTGTACTCCTGAACAATGGAAACAACCTTAGAGACAGAAATAAACAACACTTCGCTTATGTA	6240
Qy	6241	GGACCTCTTGTTCCTCAATSTGAAATACCCATACCTCAAAACCTCACCTGTGTAAATTT	6300
Db	6241	GGACCTCTTGTTCCTCAATSTGAAATACCCATACCTCAAAACCTCACCTGTGTAAATTT	6300
Qy	6301	AGCAATACTACATACACAACAACCTCCCAATGCAATGAGTGGGTAATCTCTCCACACAA	6360
Db	6301	AGCAATACTACATACACAACAACCTCCCAATGCAATGAGTGGGTAATCTCTCCACACAA	6360
Qy	6361	ATTAGTGCAGTAAACCTCAGGAAATATTTTGTCTGTGTAACCTCAGGCTATGCTGTGTTG	6420
Db	6361	ATTAGTGCAGTAAACCTCAGGAAATATTTTGTCTGTGTAACCTCAGGCTATGCTGTGTTG	6420
Qy	6421	AATGGCTCTTCAGAAATCTATGTGCTCTCTCAATCTTAAATGCCCCATATGRCATCTAC	6480
Db	6421	AATGGCTCTTCAGAAATCTATGTGCTCTCTCAATCTTAAATGCCCCATATGRCATCTAC	6480
Qy	6481	ACTGAAACAAGATTTATACAGTATATGATCATATCTTAAGCCCCCGCAACAAGATGCCATT	6540
Db	6481	ACTGAAACAAGATTTATACAGTATATGATCATATCTTAAGCCCCCGCAACAAGATGCCATT	6540



```

QY 6541 CTTCTTTTGTATAGAGAGAGAGTGTAGTGCATAGTACTGGCATTTGGCGGTATC 6600
DB 6541 CTTCTTTTGTATAGAGAGAGAGTGTAGTGCATAGTACTGGCATTTGGCGGTATC 6600
QY 6601 ACACCTCTACTAGTCTTCTACTACAACTATCTCAAGAACTAAATGGGACATGACAGG 6660
DB 6601 ACACCTCTACTAGTCTTCTACTACAACTATCTCAAGAACTAAATGGGACATGACAGG 6660
QY 6661 GTCCGCGATCTCTGGTCACTTGCAGAGATCACTTAAGTCTCTAGAGAGATGCTCTT 6720
DB 6661 GTCCGCGATCTCTGGTCACTTGCAGAGATCACTTAAGTCTCTAGAGAGATGCTCTT 6720
QY 6721 CCAAAATCGAAGAGCTTATAGTCTGTACACGCTGAGAGAGGGGAGAACTGTTATTTTA 6780
DB 6721 CCAAAATCGAAGAGCTTATAGTCTGTACACGCTGAGAGAGGGGAGAACTGTTATTTTA 6780
QY 6781 GGGGAGAAATGCTGTTATATGTTAAATCAATCCGAATCGTCACTGAGAAAGTTBAAGA 6840
DB 6781 GGGGAGAAATGCTGTTATATGTTAAATCAATCCGAATCGTCACTGAGAAAGTTBAAGA 6840
QY 6841 ATTGAGATCGAATACAGTAKAGAGAGAGCTTCAAAACATGAGACCTGGGGCTTC 6900
DB 6841 ATTGAGATCGAATACAGTAKAGAGAGAGCTTCAAAACATGAGACCTGGGGCTTC 6900
QY 6901 CTCAGCCATGAGATGCTGAGATTCCTCCCTCTTGAAGACCTGACAGCTATATATTTG 6960
DB 6901 CTCAGCCATGAGATGCTGAGATTCCTCCCTCTTGAAGACCTGACAGCTATATATTTG 6960
QY 6961 CTACTCTCTTGTGACCTCTGATCTTTTAACTCTCTGTTAACTTTGCTCTTCAAGATC 7020
DB 6961 CTACTCTCTTGTGACCTCTGATCTTTTAACTCTCTGTTAACTTTGCTCTTCAAGATC 7020
QY 7021 GAAGCTGTAAATCTACAAATGAGAGCCCAAGATGACCTCAAGCTAAAGTCTACCGCAGA 7080
DB 7021 GAAGCTGTAAATCTACAAATGAGAGCCCAAGATGACCTCAAGCTAAAGTCTACCGCAGA 7080
QY 7081 CCCCTGAGACGGGCTGTAGCCCAAGATGTTAAATGACATCAAAAGGACCCCTCTCT 7140
DB 7081 CCCCTGAGACGGGCTGTAGCCCAAGATGTTAAATGACATCAAAAGGACCCCTCTCT 7140
QY 7141 GAGGAATCTGAGCTGACCAACCTTACTACGCCCCCAATTCAGAGAGAGAGATGAGAC 7200
DB 7141 GAGGAATCTGAGCTGACCAACCTTACTACGCCCCCAATTCAGAGAGAGAGATGAGAC 7200
QY 7201 GGTGCTGGGCACTCTCCCAACAGACTTAAGTCTCTGTTGAGATGGGAGCTGAGA 7260
DB 7201 GGTGCTGGGCACTCTCCCAACAGACTTAAGTCTCTGTTGAGATGGGAGCTGAGA 7260
QY 7261 GACAGAGCTAGCTGATTTCTTAGGCTGATTAAGATCCYTAAGCTTAGTGGGAGAGTG 7320
DB 7261 GACAGAGCTAGCTGATTTCTTAGGCTGATTAAGATCCYTAAGCTTAGTGGGAGAGTG 7320
QY 7321 ACCACATCCACTTTAAACACGAGGCTTGCACTTAGTCAACCTGACCAATCAGAGAG 7380
DB 7321 ACCACATCCACTTTAAACACGAGGCTTGCACTTAGTCAACCTGACCAATCAGAGAG 7380
QY 7381 CTACATTAATAGCTAATAGGCAAGAGAGAGTAAAGAAATAGCCATCTATTATTCG 7440
DB 7381 CTACATTAATAGCTAATAGGCAAGAGAGAGTAAAGAAATAGCCATCTATTATTCG 7440
QY 7441 MTGAGAGAGAGAGAGAGAGAGCAATGATGGGATTTAAACCAAGTTCGAGCGGAGAA 7500
DB 7441 MTGAGAGAGAGAGAGAGAGAGCAATGATGGGATTTAAACCAAGTTCGAGCGGAGAA 7500
QY 7501 CGGCAACCCCTTTGGGCTCCCTCTCTTGTATGGAGCTCTGTTTCACTATTTTCA 7560
DB 7501 CGGCAACCCCTTTGGGCTCCCTCTCTTGTATGGAGCTCTGTTTCACTATTTTCA 7560
QY 7561 TCTATTAATCTTGACATGCR 7582
DB 7561 TCTATTAATCTTGACATGCR 7582

```

```

RESULT 2
BD267487
LOCUS
DEFINITION
  BD267487 7582 bp DNA linear PAT 17-JUL-2003
  Endogenetic nucleic fragment associated with an autoimmune disease,
  marking method and reagent.
ACCESSION
  BD267487
  BD267487.1 GI:33077255
VERSION
  UP 2002534980-A/28.
KEYWORDS
  Homo sapiens (human)
SOURCE
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
  1 (bases 1 to 7582)
  Baccala, G.P., Mallet, F. and Voisset, C.
  Endogenetic nucleic fragment associated with an autoimmune disease,
  marking method and reagent
  Patent: JP 2002534980-A 28 22-OCT-2002;
  BIO MERIEUX
JOURNAL
  OS Homo sapiens (human)
  PN JP 2002534980-A/28
  BD 22-OCT-2002
  PF 21-JAN-2000 JP 2000594929
  PI 21-JAN-1999 FR 99/00888
  PI GLAUDIA PARANHOS BACCALA, FRANCOIS MALLET, CECILE VOISSET PC
  C12N15/09, C07K14/15, C12Q1/02, C12Q1/68, G01N33/53, G01N33/53, PC
  G01N33/564
  PC G01N33/566, G01N33/569, G01N33/58//C12P21/02, C12N15/00 CC
  Endogenetic nucleic fragment associated with an autoimmune CC
  disease,
  CC marking method and reagent
  FH Key Location/Qualifiers
  FT source 1.7582
  FT 1.7582 /organism="Homo sapiens (human)".
  1.7582 Location/Qualifiers
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
FEATURES
  source
  ORIGIN
  Query Match 98.8%; Score 7489.6; DB 6; Length 7582;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1 CAACATGGGATATTAACCCAGGCAATTCGAGCTG6CAACAGACAGCCCCCTTTGGTCC 60
  DB 1 CAACATGGGATATTAACCCAGGCAATTCGAGCTG6CAACAGACAGCCCCCTTTGGTCC 60
  QY 61 CTCCCTTTGTATGAGAGCTGTTTCAATGCTATTCACTTAAATCTTGCAACTGCA 120
  DB 61 CTCCCTTTGTATGAGAGCTGTTTCAATGCTATTCACTTAAATCTTGCAACTGCA 120
  QY 121 CTCTTGCTGATGATTTCTTAACGCTGAGCTGAGCTTTGCTCAACGCTCAACACTGC 180
  DB 121 CTCTTGCTGATGATTTCTTAACGCTGAGCTGAGCTTTGCTCAACGCTCAACACTGC 180
  QY 121 CTCTTGCTGATGATTTCTTAACGCTGAGCTGAGCTTTGCTCAACGCTCAACACTGC 180
  DB 121 CTCTTGCTGATGATTTCTTAACGCTGAGCTGAGCTTTGCTCAACGCTCAACACTGC 180
  QY 181 TGTTCGCAACACCGCANAACTCGCGCTGACTCCCATCCCTCTGATCTCTGAGAGGTGC 240
  DB 181 TGTTCGCAACACCGCANAACTCGCGCTGACTCCCATCCCTCTGATCTCTGAGAGGTGC 240
  QY 241 CGCTGTGCTCTGATTCAGAGAGAGGAGCCATTGCGCGCTCCCAATTGGGCTAAAGCTTC 300
  DB 241 CGCTGTGCTCTGATTCAGAGAGAGGAGCCATTGCGCGCTCCCAATTGGGCTAAAGCTTC 300
  QY 301 CATGTCCTGACAGGCTTAAGTGCCTGGGTTTGTCTAATGAGCTGAACATANTCACT 360
  DB 301 CATGTCCTGACAGGCTTAAGTGCCTGGGTTTGTCTAATGAGCTGAACATANTCACT 360
  QY 361 GGGTTCATGCTTCTTCTGTGACCAAGGCTTCTAATAAATTAACAATTACCA 420
  DB 361 GGGTTCATGCTTCTTCTGTGACCAAGGCTTCTAATAAATTAACAATTACCA 420

```

OY	421	TGCCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAGACTCCAGGTCAGAGAAATTC	480
Db	421	TGGCCCAAGATTCCATTCCTTGGAATCCGTGAGGSCAAGACTCCAGGTCAGAGAAATTC	480
OY	481	GARGCTTGGCCACCAATCTTGGAAAGGGGCGTGACCTCTTGGAAAGTGGTTCACCAACATC	540
Db	481	GARGCTTGGCCACCAATCTTGGAAAGGGGCGTGACCTCTTGGAAAGTGGTTCACCAACATC	540
OY	541	TTGGAGAGCTCTGTGAGCAAGACCCCGGTTCACATTTTGGCRAACAMSPACGGACATCC	600
Db	541	TTGGAGAGCTCTGTGAGCAAGACCCCGGTTCACATTTTGGCRAACAMSPACGGACATCC	600
OY	601	MAAGTATGGGAAAAGTTCCTCCCGCAAGCAAAAAAGCCCTTAAGAGGTATTTCTGGARAT	660
Db	601	MAAGTATGGGAAAAGTTCCTCCCGCAAGCAAAAAAGCCCTTAAGAGGTATTTCTGGARAT	660
OY	661	TGGGAMCAAATTTGACCCCTCAGACACTAAGAAAGAAAGACCTTATCTTCTGCAAGTGC	720
Db	661	TGGGAMCAAATTTGACCCCTCAGACACTAAGAAAGAAAGACCTTATCTTCTGCAAGTGC	720
OY	721	GCCTGCACTCCTGAGGGAGATTAATAATTATACACCATCTTACAGCTAGACCTTCTTGG	780
Db	721	GCCTGCACTCCTGAGGGAGATTAATAATTATACACCATCTTACAGCTAGACCTTCTTGG	780
OY	781	TAGAAAAGGCAATGGAGTGAAGTGCATAGTACAACTTCTCTTTCAATTAAAGACAA	840
Db	781	TAGAAAAGGCAATGGAGTGAAGTGCATAGTACAACTTCTCTTTCAATTAAAGACAA	840
OY	841	CTCACAAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTAGAGTCACTCTCC	900
Db	841	CTCACAAATTATGTAAAAAGTGTGATTTATGCCCTACAGGAAGCCTTAGAGTCACTCTCC	900
OY	901	CTATCCCAAGCATCCCCGACCTCCCTCCCAATTAATAAGAACCCCTTCAACCCAAATGG	960
Db	901	CTATCCCAAGCATCCCCGACCTCCCTCCCAATTAATAAGAACCCCTTCAACCCAAATGG	960
OY	961	TCCAAAAGGAGATAGACAAAGAGGTAAACAGTAAACAAAGATGCCAATATTTCCCAAT	1020
Db	961	TCCAAAAGGAGATAGACAAAGAGGTAAACAGTAAACAAAGATGCCAATATTTCCCAAT	1020
OY	1021	TATGACCCCTCCCAAGCACTGGAGAGAAAGAAATTCGGCCGACAGGCAAGTGCATGTGCT	1080
Db	1021	TATGACCCCTCCCAAGCACTGGAGAGAAAGAAATTCGGCCGACAGGCAAGTGCATGTGCT	1080
OY	1081	TTTTTCTCCGACACTTAAGCAATTAATAACAGACTAGGTAAATTTCTCAGTAATACCT	1140
Db	1081	TTTTTCTCCGACACTTAAGCAATTAATAACAGACTAGGTAAATTTCTCAGTAATACCT	1140
OY	1141	GATGGCTATATTTGATGTTTTCAGAGGTTTAGCAATTTCTTTGATCTGACATGAGAGAT	1200
Db	1141	GATGGCTATATTTGATGTTTTCAGAGGTTTAGCAATTTCTTTGATCTGACATGAGAGAT	1200
OY	1201	ATATATGTCACTGTCTAATTCACACACTAACCCCAATAGAGAGGTGCACATTAATGCG	1260
Db	1201	ATATATGTCACTGTCTAATTCACACACTAACCCCAATAGAGAGGTGCACATTAATGCG	1260
OY	1261	AGCCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGAGTANGATGACACA	1320
Db	1261	AGCCTGAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGAGTANGATGACACA	1320
OY	1321	GAAAGAAAGANAATGATTTCCCAAGGCCAGCAGCAGAGTCCCACTTASACCTCATTG	1380
Db	1321	GAAAGAAAGANAATGATTTCCCAAGGCCAGCAGCAGAGTCCCACTTASACCTCATTG	1380
OY	1381	GGGACACAGAAATCACTAACATGGAGATTTGGGTGCGACATTTGCTAACTTGTGTGC	1440
Db	1381	GGGACACAGAAATCACTAACATGGAGATTTGGGTGCGACATTTGCTAACTTGTGTGC	1440
OY	1441	TASAAAGACTAAGGAAAATASAGAAAAARTCTAAGATTAATCTCAATGATGTCCACATA	1500
Db	1441	TASAAAGACTAAGGAAAATASAGAAAAARTCTAAGATTAATCTCAATGATGTCCACATA	1500
OY	1501	ACACAGGGAGGGAAGAAAATCTTACTGCTTTTCTGAGAGACTAAGGAGCAATTGAG	1560

Db	1501	ACACAGGGGAAAGGAAAGAAATCTACCTCTTTCTGGAGACCTAAGGAGGCATTGG	1560
Qy	1561	GAAAGCGGCTCTCTGCTCACTGACTCTTCTGAAAGCCAACTAATCTTAAAGCGTAAGTT	1620
Db	1561	GAAGCGGCTCTCTGCTCACTGACTCTTCTGAAAGCCAACTAATCTTAAAGCGTAAGTT	1620
Qy	1621	TATCACTCACTCACTGACAGACATTAGAAAAAACTTCAAAAGCTGCGTAGGCCCGGAG	1680
Db	1621	TATCACTCACTCACTGACAGACATTAGAAAAAACTTCAAAAGCTGCGTAGGCCCGGAG	1680
Qy	1681	CAAAACTTAAGAAACCCCTAATTGAACCTTGGGAACTGCGTTTTTATATAGAGATCAAGAG	1740
Db	1681	CAAAACTTAAGAAACCCCTAATTGAACCTTGGGAACTGCGTTTTTATATAGAGATCAAGAG	1740
Qy	1741	GAGCAGGCGGAAACAGGACAAACGGAGTTAAAAAAAAGGCCACCGCTTTAGTCATGACCT	1800
Db	1741	GAGCAGGCGGAAACAGGACAAACGGAGTTAAAAAAAAGGCCACCGCTTTAGTCATGACCT	1800
Qy	1801	CAGGCAAGTGAAGCTTTGGAAGCTCTGGAAGAAAGGAAAGCTGGGCAATTGAATGCTTAA	1860
Db	1801	CAGGCAAGTGAAGCTTTGGAAGCTCTGGAAGAAAGGAAAGCTGGGCAATTGAATGCTTAA	1860
Qy	1861	TAGGGCTTGCTCTCACTGCGGCTTCAAGGACCTTAAAGAAATGTCGCAAGTGAAG	1920
Db	1861	TAGGGCTTGCTCTCACTGCGGCTTCAAGGACCTTAAAGAAATGTCGCAAGTGAAG	1920
Qy	1921	TAAACCGCCCTTCGTCATAGCCCTTAATTCAAGGAAATCACTGAAAGGCCCACTGCC	1980
Db	1921	TAAACCGCCCTTCGTCATAGCCCTTAATTCAAGGAAATCACTGAAAGGCCCACTGCC	1980
Qy	1981	CAGGGGACAAAGGCTTTTGAAGTCAGAGGCCACTAACCAGATATCCAGCAGCAGACTG	2040
Db	1981	CAGGGGACAAAGGCTTTTGAAGTCAGAGGCCACTAACCAGATATCCAGCAGCAGACTG	2040
Qy	2041	AGGGTCCCTGGGGCAAGCCCATTCATGECATCACCCCTCAAGAGCCCTGGGTAAGCTT	2100
Db	2041	AGGGTCCCTGGGGCAAGCCCATTCATGECATCACCCCTCAAGAGCCCTGGGTAAGCTT	2100
Qy	2101	GACCATGAGGGGCAAGAGGGTGTCTCCAGGACACTGAGTGCAGCTCTTAACTCTTACT	2160
Db	2101	GACCATGAGGGGCAAGAGGGTGTCTCCAGGACACTGAGTGCAGCTCTTAACTCTTACT	2160
Qy	2161	CTTCTGTCCCGGACAACTGTCTTCAGATCTGTCACTATCTGAGGGGTCNTAAGACG	2220
Db	2161	CTTCTGTCCCGGACAACTGTCTTCAGATCTGTCACTATCTGAGGGGTCNTAAGACG	2220
Qy	2221	GCGAGTCACTAGATCTTTTCCAGGCACTAAGTTATGAACGGGGAGCTTATCTTT	2280
Db	2221	GCGAGTCACTAGATCTTTTCCAGGCACTAAGTTATGAACGGGGAGCTTATCTTT	2280
Qy	2281	TCACTATGCTTTTCTAATATATGCTTGAAGCCGCCACTGCTTAGGGAGAGCACTCTA	2340
Db	2281	TCACTATGCTTTTCTAATATATGCTTGAAGCCGCCACTGCTTAGGGAGAGCACTCTA	2340
Qy	2341	GCAAAAGCAGGGGCCATTTATACACTGTACATAGAGAAAGAACACCCGTTTGTGTCNC	2400
Db	2341	GCAAAAGCAGGGGCCATTTATACACTGTACATAGAGAAAGAACACCCGTTTGTGTCNC	2400
Qy	2401	CTGCTTGAAGAAAGAAATTAATCTCTGAAGTTTGGGCAACGAAGACAAATATGAGCAGC	2460
Db	2401	CTGCTTGAAGAAAGAAATTAATCTCTGAAGTTTGGGCAACGAAGACAAATATGAGCAGC	2460
Qy	2461	CAAAAGAAAGCCGCTCTGCTTCAATTAATCTAAGGATTCACCTTCCTTCCATACAAA	2520
Db	2461	CAAAAGAAAGCCGCTCTGCTTCAATTAATCTAAGGATTCACCTTCCTTCCATACAAA	2520
Qy	2521	GCGAGTACCCCTCAGACCCAAAGGCCCAAGAGTTCCAAAAGATGTTAAGGACTTAA	2580
Db	2521	GCGAGTACCCCTCAGACCCAAAGGCCCAAGAGTTCCAAAAGATGTTAAGGACTTAA	2580
Qy	2581	AAGCCAGGCTTAAATACATGCAATATCTCCCTGAGATATTCGTTAGATTGAG	2640
Db	2581	AAGCCAGGCTTAAATACATGCAATATCTCCCTGAGATATTCGTTAGATTGAG	2640

Db 2581 AAGCCCAAGCTTAGTAAAAACATGATACCTCCCTGCAATATTCCTGATGATGAG 2640  
Qy 2641 GAGCACAGAAACCAGTGTGACAGTGGAGGTTAGTGCAGATCTCAGATTTATCAATGG 2700  
Db 2641 GAGGCAGAGAAACCAAGTGACAGTGGAGGTTAGTGCAGATCTCAGATTTATCAATGG 2700  
Qy 2701 AGGCGTGTCTCTTTTATACCAGCTGTACTAGCCCTTATATCTGTGMYTTCGAAATAC 2760  
Db 2701 AGGCGTGTCTCTTTTATACCAGCTGTACTAGCCCTTATATCTGTGMYTTCGAAATAC 2760  
Qy 2761 CAGAGGAAGAGATGGTTTACSTCCGAGACCTTMAAGATGCTTCTTCTGCATCCCTG 2820  
Db 2761 CAGAGGAAGAGATGGTTTACSTCCGAGACCTTMAAGATGCTTCTTCTGCATCCCTG 2820  
Qy 2821 TACATCTGACTCTCAATTCCTGTTGCTTGAAGATCTTCAAAACCCATCTCAAC 2880  
Db 2821 TACATCTGACTCTCAATTCCTGTTGCTTGAAGATCTTCAAAACCCATCTCAAC 2880  
Qy 2881 TCACCTGACTRTTTTACCCCAAGGTTCAGGATAGYCCCATCTATTTGGCAGGCAAT 2940  
Db 2881 TCACCTGACTRTTTTACCCCAAGGTTCAGGATAGYCCCATCTATTTGGCAGGCAAT 2940  
Qy 2941 TAGCCCAAGACTGAGCATATMTCATACCTGACATCTTCTCTCTGATGATG 3000  
Db 2941 TAGCCCAAGACTGAGCATATMTCATACCTGACATCTTCTCTGATGATG 3000  
Qy 3001 ATTACTTTTGGGCGCTTCAGAAAACCTTGTGCATCAAGCCCAAGACCTCTTMA 3060  
Db 3001 ATTACTTTTGGGCGCTTCAGAAAACCTTGTGCATCAAGCCCAAGACCTCTTMA 3060  
Qy 3061 ATTTCTCTGCGYACCTGTGCTCAWGGTTTCCAAAGSABARBCCTCATCTGCTCAGAC 3120  
Db 3061 ATTTCTCTGCGYACCTGTGCTCAWGGTTTCCAAAGSABARBCCTCATCTGCTCAGAC 3120  
Qy 3121 AGGTTAAATCTTAAGBCTAABATTAATCCAAAGGACACARGGCCCTCAGTGAAGAAATTA 3180  
Db 3121 AGGTTAAATCTTAAGBCTAABATTAATCCAAAGGACACARGGCCCTCAGTGAAGAAATTA 3180  
Qy 3181 TCCAGGCTATATCTGCTTATCTCATCTCATCAAAACCTTAAGCAACTAAGBRRTTCTTG 3240  
Db 3181 TCCAGGCTATATCTGCTTATCTCATCTCATCAAAACCTTAAGCAACTAAGBRRTTCTTG 3240  
Qy 3241 GCTTAAYAGGYTCTGCGAAMAATGATCCCGAGTWTGCRPAATAGCCAGGYCATTA 3300  
Db 3241 GCTTAAYAGGYTCTGCGAAMAATGATCCCGAGTWTGCRPAATAGCCAGGYCATTA 3300  
Qy 3301 WATACASTAATTAAAGAACTCAGAAAGCCAAATACCATTATTAATGAATGAATMCTGAA 3360  
Db 3301 WATACASTAATTAAAGAACTCAGAAAGCCAAATACCATTATTAATGAATGAATMCTGAA 3360  
Qy 3361 GYWRBAAGTGGCTTCCAGGCCCCCTTAAAGAGGCTTAAACCAAGYCCGAGTGTAAAT 3420  
Db 3361 GYWRBAAGTGGCTTCCAGGCCCCCTTAAAGAGGCTTAAACCAAGYCCGAGTGTAAAT 3420  
Qy 3421 TGGCAACRGGGCAAGACTTTTSTTATATATCAAGAAAAAAACAGAAAYAGCTCTRGGA 3480  
Db 3421 TGGCAACRGGGCAAGACTTTTSTTATATATCAAGAAAAAAACAGAAAYAGCTCTRGGA 3480  
Qy 3481 GTTCCTTACACAGTTCGAGGAGATGAGTTGCAACCTRTGGCRATCTGASTAAGAAAT 3540  
Db 3481 GTTCCTTACACAGTTCGAGGAGATGAGTTGCAACCTRTGGCRATCTGASTAAGAAAT 3540  
Qy 3541 GATGTAGTGGCAAAAGGTTGCTTATTTTAAAGGATGAGTGGGAGAGTGAAGAGTGTAA 3600  
Db 3541 GATGTAGTGGCAAAAGGTTGCTTATTTTAAAGGATGAGTGGGAGAGTGAAGAGTGTAA 3600  
Qy 3601 GTATCTGAAGCACTTAAATATATACAGGAGAGATCTTACTGTGAGACATCTCATGAK 3660  
Db 3601 GTATCTGAAGCACTTAAATATATACAGGAGAGATCTTACTGTGAGACATCTCATGAK 3660  
Qy 3661 GTGAATRGATATCTCACTGTCTAAAGAGACTTGTGGCTGTCAGACAACTTTACTTAA 3720  
Db 3661 GTGAATRGATATCTCACTGTCTAAAGAGACTTGTGGCTGTCAGACAACTTTACTTAA 3720

Qy 3721 TPTCAGGCTCTATTAATCTTAAGGCGCAGTGTGACATCTGACATCTTGCAACTCTTAAC 3780  
Db 3721 TPTCAGGCTCTATTAATCTTAAGGCGCAGTGTGACATCTGACATCTTGCAACTCTTAAC 3780  
Qy 3781 CCAGYCNCAATTTCTTCAGACAAATGAAGAAAGTAAATATATCTGCAACATTAAT 3840  
Db 3781 CCAGYCNCAATTTCTTCAGACAAATGAAGAAAGTAAATATATCTGCAACATTAAT 3840  
Qy 3841 TCTCAAACTATGCAAGGAGCTTATGAGTTCCTTGAATCTGATTCYGAACCT 3900  
Db 3841 TCTCAAACTATGCAAGGAGCTTATGAGTTCCTTGAATCTGATTCYGAACCT 3900  
Qy 3901 CAACTGTATATGATGAAGAGTTCCTTGTGAAGAAAGAGCTTGAAAGYGGGATGAC 3960  
Db 3901 CAACTGTATATGATGAAGAGTTCCTTGTGAAGAAAGAGCTTGAAAGYGGGATGAC 3960  
Qy 3961 AGTGTCACTGATTAATGAATATATGAAATGAAATGCTTCACTCAGGAACTAGTCTTA 4020  
Db 3961 AGTGTCACTGATTAATGAATATATGAAATGAAATGCTTCACTCAGGAACTAGTCTTA 4020  
Qy 4021 GCTRGCAAACTAATAGCCCTCATYTTGGGCACTAGAAATTTAGAGAAAGGATTA 4080  
Db 4021 GCTRGCAAACTAATAGCCCTCATYTTGGGCACTAGAAATTTAGAGAAAGGATTA 4080  
Qy 4081 TATATATACAGACTCTAATATGCTTAACCTAAGTCTCATGCTGATGACATATATGAT 4140  
Db 4081 TATATATACAGACTCTAATATGCTTAACCTAAGTCTCATGCTGATGACATATATGAT 4140  
Qy 4141 AGAAAGGAATTCCTAATCTTGAAGGAAACCTTATCAMAATCAGAAACCAATTAAGAR 4200  
Db 4141 AGAAAGGAATTCCTAATCTTGAAGGAAACCTTATCAMAATCAGAAACCAATTAAGAR 4200  
Qy 4201 ATTAATATATGCGMGTACAAAACCTAAGAGTGAAGTCTTACATGCTGGGCTCAT 4260  
Db 4201 ATTAATATATGCGMGTACAAAACCTAAGAGTGAAGTCTTACATGCTGGGCTCAT 4260  
Qy 4261 NAAAGGAAGAAAGGAAATASAAAGGAATGSCAAGCAKATTTGAAGCMAAAGAGC 4320  
Db 4261 NAAAGGAAGAAAGGAAATASAAAGGAATGSCAAGCAKATTTGAAGCMAAAGAGC 4320  
Qy 4321 TGCAGGCAAGACCTCTCATTAAGAAATGCTTATTAACCTTCTGATATAGGATATCC 4380  
Db 4321 TGCAGGCAAGACCTCTCATTAAGAAATGCTTATTAACCTTCTGATATAGGATATCC 4380  
Qy 4381 CTTCCGGGAAACCAAGCCCCCACTATCAGCAGAGAAACAGATGGGAACTCAGAGG 4440  
Db 4381 CTTCCGGGAAACCAAGCCCCCACTATCAGCAGAGAAACAGATGGGAACTCAGAGG 4440  
Qy 4441 CAGTTTCTCCCTCGGAGCGGTTAGCCACTGAAGAAAGGAAATACTTTTGTGCTCAAC 4500  
Db 4441 CAGTTTCTCCCTCGGAGCGGTTAGCCACTGAAGAAAGGAAATACTTTTGTGCTCAAC 4500  
Qy 4501 TATCCAAATGAAATTAATTAACCTTCAATCAACCTTCACTTGAAGCATGATGAC 4560  
Db 4501 TATCCAAATGAAATTAATTAACCTTCAATCAACCTTCACTTGAAGCATGATGAC 4560  
Qy 4561 CCATCARATGGCCAAATCTTATTAATGAGCAAGGCTTTTCAAACTATCAAGABAT 4620  
Db 4561 CCATCARATGGCCAAATCTTATTAATGAGCAAGGCTTTTCAAACTATCAAGABAT 4620  
Qy 4621 AKTCAGGGCTGTGAATGTGCAABABAAATATCCCTGCTCATGCGCAAGCTCTTC 4680  
Db 4621 AKTCAGGGCTGTGAATGTGCAABABAAATATCCCTGCTCATGCGCAAGCTCTTC 4680  
Qy 4681 AGGABAAACAABAAACAAGGCAATTAACCTGABABABAATGGAATTTTACCAAG 4740  
Db 4681 AGGABAAACAABAAACAAGGCAATTAACCTGABABABAATGGAATTTTACCAAG 4740  
Qy 4741 CCCAAACCTCAGAGGATTTCAAGATCTAATGCTGGGTBARAATCTTCAAGGCTTGGA 4800  
Db 4741 CCCAAACCTCAGAGGATTTCAAGATCTAATGCTGGGTBARAATCTTCAAGGCTTGGA 4800



Db 6961 CTACTCTCTTGGACCCCTGTATCTTACCTCCTTGTAACTTGTGCTCTTCCAGATC 7020  
Qy 7021 GAAGCTGTAAACTACAAATGAGAGCCCAAGATGAGTCCAGACCTTAAGATGTACCGAGA 7080  
Db 7021 GAAGCTGTAAACTACAAATGAGAGCCCAAGATGAGTCCAGACCTTAAGATGTACCGAGA 7080  
Qy 7081 CCCCTGAGCCGGCTGTAGCCCAAGATGAGTCCAGACCTTAAGATGTACCGAGA 7140  
Db 7081 CCCCTGAGCCGGCTGTAGCCCAAGATGAGTCCAGACCTTAAGATGTACCGAGA 7140  
Qy 7141 GAGGAATCTCAGTGTGACAACTCTTCTAGAGCCCAATGAGAGAGAGAGAGAGAGAG 7200  
Db 7141 GAGGAATCTCAGTGTGACAACTCTTCTAGAGCCCAATGAGAGAGAGAGAGAGAGAG 7200  
Qy 7201 GGTGTGCGGCAACCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7260  
Db 7201 GGTGTGCGGCAACCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7260  
Qy 7261 GACAGAGCTAGCTGATTTCTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7320  
Db 7261 GACAGAGCTAGCTGATTTCTTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7320  
Qy 7321 ACCACATCCACCTTTTAAACAGGGGCTTGCACCTTATGACACCTTACCAATCAGAG 7380  
Db 7321 ACCACATCCACCTTTTAAACAGGGGCTTGCACCTTATGACACCTTACCAATCAGAG 7380  
Qy 7381 CTCCTAAATGCTAAATTAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7440  
Db 7381 CTCCTAAATGCTAAATTAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7440  
Qy 7441 MTGAGAGCAAG 7500  
Db 7441 MTGAGAGCAAG 7500  
Qy 7501 CGGCAACCCCTTTGGGTCCCTCTTGTATGAGAGCTCTTTTATGCTATTTTAC 7560  
Db 7501 CGGCAACCCCTTTGGGTCCCTCTTGTATGAGAGCTCTTTTATGCTATTTTAC 7560  
Qy 7561 TCTATTAATCTTGACACTGCR 7582  
Db 7561 TCTATTAATCTTGACACTGCR 7582

RESULT 3  
AX000966 7582 bp DNA 1 linear PAT 10-MAR-2000  
LOCUS AX000966  
DEFINITION Sequence 11 from Patent WO9902696.  
ACCESSION AX000966  
VERSION AX000966.1 GI:7241208  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 7582)  
AUTHORS Beseme, F. and Blond, J.  
TITLE ENDOGENETIC RETROVIRAL SEQUENCES, ASSOCIATED WITH AUTOIMMUNE  
DISEASES OR WITH PREGNANCY DISORDERS  
JOURNAL Patent: WO 9902696-A 11 21-JAN-1999;  
BIO MEDICAL (FR); BESEME, FREDERIC (FR)  
FEATURES  
source 1..7582  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3264"

Query Match 98.8%; Score 7489.6; DB 6; Length 7582;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAACAATGGGATATTAACCCAGAGCTTGGAGTGCACAGAGCCCTTTGGGTCC 60  
Db 1 CAACAATGGGATATTAACCCAGAGCTTGGAGTGCACAGAGCCCTTTGGGTCC 60

Qy 61 CTTCCTTTGATGAGAGCTTTTGAAGTATTTTCACTTATTAATCTTGAAGTCA 120  
Db 61 CTTCCTTTGATGAGAGCTTTTGAAGTATTTTCACTTATTAATCTTGAAGTCA 120  
Qy 121 CTCTCTGATGATGATTTCTTAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
Db 121 CTCTCTGATGATGATTTCTTAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
Qy 181 TGTTCGACACACCGGAACTTCCGCTGATCCCATCTTCTGATCTTGAAGTCTGAGAGG 240  
Db 181 TGTTCGACACACCGGAACTTCCGCTGATCCCATCTTCTGATCTTGAAGTCTGAGAGG 240  
Qy 241 CGCTGTCTCTGATGATGAG 300  
Db 241 CGCTGTCTCTGATGATGAG 300  
Qy 301 CATTGTCCTGACAGCTAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
Db 301 CATTGTCCTGACAGCTAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
Qy 361 GGGTTCATGATGATTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
Db 361 GGGTTCATGATGATTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
Qy 421 TGAGCCCAAGATTCATTCTTGGAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
Db 421 TGAGCCCAAGATTCATTCTTGGAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480  
Qy 481 GAGCTTGCACCATCTTGAAG 540  
Db 481 GAGCTTGCACCATCTTGAAG 540  
Qy 541 TTGGAGAGCTCTGTGAG 600  
Db 541 TTGGAGAGCTCTGTGAG 600  
Qy 601 MAAGTATGAG 660  
Db 601 MAAGTATGAG 660  
Qy 661 TGAGAGCAATTTGACCTTCAAGACCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Db 661 TGAGAGCAATTTGACCTTCAAGACCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720  
Qy 721 GCTTGGACATCTGAG 780  
Db 721 GCTTGGACATCTGAG 780  
Qy 781 TAGAAGAGCAATGAG 840  
Db 781 TAGAAGAGCAATGAG 840  
Qy 841 CTCACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900  
Db 841 CTCACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900  
Qy 901 CTATCCAGAGATCCCGGACTCTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAAT 960  
Db 901 CTATCCAGAGATCCCGGACTCTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAAT 960  
Qy 961 TCCAAAGAGATGAG 1020  
Db 961 TCCAAAGAGATGAG 1020  
Qy 1021 TATGACCCCTCCCAAG 1080  
Db 1021 TATGACCCCTCCCAAG 1080  
Qy 1081 TTTTCTTCCCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140  
Db 1081 TTTTCTTCCCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140



OY	1141	GATGCTATATTGRTGTTTACAGAGGTTAGACAACTTTTGATCTGACATGGAGAT	1200
DB	1141	GATGGCTATATGRTGTTTACAGAGGTTAGACAACTTTTGATCTGACATGGAGAT	1200
OY	1201	ATAATGTCACTGCTAAATCAGACACTAACCCCAATAGAGAAATGCACATACTGC	1260
DB	1201	ATAATGTCACTGCTAAATCAGACACTAACCCCAATAGAGAAATGCACATACTGC	1260
OY	1261	AGCCTGAGRGTTCGCGATCTCTGGTATCTCAGTCAGGTCAATGATATAGATACACA	1320
DB	1261	AGCCTGAGRGTTCGCGATCTCTGGTATCTCAGTCAGGTCAATGATATAGATACACA	1320
OY	1331	GAGGAAAGNAAATATATTTCCCAAGGCCACAGCACTTCCCACTATACCTCTATG	1380
DB	1331	GAGGAAAGNAAATATATTTCCCAAGGCCACAGCACTTCCCACTATATGCTATG	1380
OY	1391	GAGGAAAGNAAATATATTTCCCAAGGCCACAGCACTTCCCACTATACCTCTATG	1440
DB	1391	GAGGAAAGNAAATATATTTCCCAAGGCCACAGCACTTCCCACTATACCTCTATG	1440
OY	1441	TASAAAGACTAAGGAAAACTASGAAGAARTCTATGATTACTCATGATGTCACATA	1500
DB	1441	TASAAAGACTAAGGAAAACTASGAAGAARTCTATGATTACTCATGATGTCACATA	1500
OY	1501	ACACAGGGGAAAGGAAAGAAATCCTACTGCTTTTGGAGAGCTAAGGGAGCATTTAG	1560
DB	1501	ACACAGGGGAAAGGAAAGAAATCCTACTGCTTTTGGAGAGCTAAGGGAGCATTTAG	1560
OY	1561	GAAAGGTGCTCTCTGTGACCTGATCTCTTGAAGGCCATATCTTAAAGCGTAAGTT	1620
DB	1561	GAAAGGTGCTCTCTGTGACCTGATCTCTTGAAGGCCATATCTTAAAGCGTAAGTT	1620
OY	1621	TATCACTCAGTCAGCTGCAGACATTAAGAAAACTTCAAAGCTGCGGTAGGCCGAG	1680
DB	1621	TATCACTCAGTCAGCTGCAGACATTAAGAAAACTTCAAAGCTGCGGTAGGCCGAG	1680
OY	1681	CAAAACTTAGAAACCTTATTTGAATCTTGCAACCTTGCTTTTATATATGATCAGAG	1740
DB	1681	CAAAACTTAGAAACCTTATTTGAATCTTGCAACCTTGCTTTTATATATGATCAGAG	1740
OY	1741	GAGCAGGGGGAACAGACAAACGGGATTAAGAAAAAGGCCACCGTTATGATGACCT	1800
DB	1741	GAGCAGGGGGAACAGACAAACGGGATTAAGAAAAAGGCCACCGTTATGATGACCT	1800
OY	1801	CAGGCAAGTGACACTTTGGAGGCTCTGAGAAAAAGGAAAGCTGGCAATTGATGCTAA	1860
DB	1801	CAGGCAAGTGACACTTTGGAGGCTCTGAGAAAAAGGAAAGCTGGCAATTGATGCTAA	1860
OY	1861	TAGGGCTTGCTTCAGTGCGGCTTACAGAGACATTTAAGAAAAATGTTCCAGATGAG	1920
DB	1861	TAGGGCTTGCTTCAGTGCGGCTTACAGAGACATTTAAGAAAAATGTTCCAGATGAG	1920
OY	1921	TAAAGCGCCCTTGCTGCTCCATGCCCCCTTATTTAAGGAATCACTGGAAAGGCCACTGCC	1980
DB	1921	TAAAGCGCCCTTGCTGCTCCATGCCCCCTTATTTAAGGAATCACTGGAAAGGCCACTGCC	1980
OY	1981	CAGGGGAAACAAGGCTTTTGAATCAGAAAGCCATAACAGATGATCCAGACAGAGACTG	2040
DB	1981	CAGGGGAAACAAGGCTTTTGAATCAGAAAGCCATAACAGATGATCCAGACAGAGACTG	2040
OY	2041	AGGGTGCTGGGGCAAGCGCCATCCCATGCACTCAACCTTCACAGAGGCCCTGGATAGCTT	2100
DB	2041	AGGGTGCTGGGGCAAGCGCCATCCCATGCACTCAACCTTCACAGAGGCCCTGGATAGCTT	2100
OY	2101	GACCAATTGAGGCCCAGGAAGTTGTCTCTGGAACACTGGTGGGCTTCTTATGCTTAAT	2160
DB	2101	GACCAATTGAGGCCCAGGAAGTTGTCTCTGGAACACTGGTGGGCTTCTTATGCTTAAT	2160
OY	2161	CTTCTGTCGCCGGAACAATGCTCTCCAGATCTGTCACTATCTGAGAGGGGCTCANTAAAGC	2220
DB	2161	CTTCTGTCGCCGGAACAATGCTCTCCAGATCTGTCACTATCTGAGAGGGGCTCANTAAAGC	2220
OY	2221	GGCAGATCACTAGATCTTTTTCACAGCACTAAGTTATGAATGGGAGCTTATTTCTT	2280

Db	2221	GGCAGTAC	TAGTATGCTTTTCCCAAGCCATAAGTATGATCTGGGAGCTTATCTTT	2280
Qy	2281	TCAATGCTTTTCTAATTATGCTTGAAAGCCCACTACCTTGTTAGGGAGACATTTCTA		2340
Db	2281	TCACATGCTTTTCTAATTATGCTTGAAAGCCCACTACCTTGTTAGGGAGACATTTCTA		2340
Qy	2341	GCAAAAGCAGGGCCATTAACACCTGAAACATATGAGAGAGAAACACCCTGTGTGTCGNC		2400
Db	2341	GCAAAAGCAGGGCCATTAACACCTGAAACATATGAGAGAGAAACACCCTGTGTGTCGNC		2400
Qy	2401	CTGCTTGAGAGAGAAATTAATCTGAAAGCTGGGCAACAGAGACATATGAGACGAGC		2460
Db	2401	CTGCTTGAGAGAGAAATTAATCTGAAAGCTGGGCAACAGAGACATATGAGACGAGC		2460
Qy	2461	CAAGAATGCGCGTCTGTTCAAGTTAACTAAAGATTCACTTCTTTCTCTACCAA		2520
Db	2461	CAAGAATGCGCGTCTGTTCAAGTTAACTAAAGATTCACTTCTTTCTCTACCAA		2520
Qy	2521	GGCAGTACCCCTTCAGACCCCAAGGCCCAACAAGATTCCAAAAGTTGTTAAGACCTTA		2580
Db	2521	GGCAGTACCCCTTCAGACCCCAAGGCCCAACAAGATTGTTAAGACCTTA		2580
Qy	2581	AAGCCCAAGGCTAGTAAACATGACATCTCCCTGACATTAATCCGTAGTGATAG		2640
Db	2581	AAGCCCAAGGCTAGTAAACATGACATCTCCCTGACATTAATCCGTAGTGATAG		2640
Qy	2641	GAGGCACAGAAACCCAGTAGACAGTGGAGGTTAGTCAAGATCTCAGATTATCAATGG		2700
Db	2641	GAGGCACAGAAACCCAGTAGACAGTGGAGGTTAGTCAAGATCTCAGATTATCAATGG		2700
Qy	2701	AGGCGCTTGCTTTTATACCCAGCTGTACTTACCTTATACGTGTGTTTTCCAAATAC		2760
Db	2701	AGGCGCTTGCTTTTATACCCAGCTGTACTTACCTTATACGTGTGTTTTCCAAATAC		2760
Qy	2761	CAGAGGAAGACAGTGGTTTACA	STCTGTGACCTTMAAGATGCGCTTCTTCTGACATCCCTG	2820
Db	2761	CAGAGGAAGACAGTGGTTTACA	STCTGTGACCTTMAAGATGCGCTTCTTCTGACATCCCTG	2820
Qy	2821	TACATCTGACTCTCAATTTCTTGTGTCCTTGAAGTACTTCAACCCACATCTCAAC		2880
Db	2821	TACATCTGACTCTCAATTTCTTGTGTCCTTGAAGTACTTCAACCCACATCTCAAC		2880
Qy	2881	TCACCTGGACTRTTATACCCCAAGGGTTACAGGGATAGTCCCACTATTTGGCCAGGAT		2940
Db	2881	TCACCTGGACTRTTATACCCCAAGGGTTACAGGGATAGTCCCACTATTTGGCCAGGAT		2940
Qy	2941	TAGCCCAAGCTGAGTCAATWTATACACCTGACATCTGTGCTTCRGTAKGTGATG		3000
Db	2941	TAGCCCAAGCTGAGTCAATWTATACACCTGACATCTGTGCTTCRGTAKGTGATG		3000
Qy	3001	ATTATCTTTTGGCGCYRTTCAGAAAACCTGTGTCATCAAGCCACCCAGC		3060
Db	3001	ATTATCTTTTGGCGCYRTTCAGAAAACCTGTGTCATCAAGCCACCCAGC		3060
Qy	3061	ATTTCCTGCGACCTGTGGCTACAMGGTTTCCAAAGSABARAGCTCARCTGCTCAACG		3120
Db	3061	ATTTCCTGCGACCTGTGGCTACAMGGTTTCCAAAGSABARAGCTCARCTGCTCAACG		3120
Qy	3121	AGGTTAAATCTTAGGCTAABATTAACAAAGGACCAAGGCCCTCAGTAGAGAAATYA		3180
Db	3121	AGGTTAAATCTTAGGCTAABATTAACAAAGGACCAAGGCCCTCAGTAGAGAAATYA		3180
Qy	3181	TCCAGGCTATAGTGGCTTATCTTCATCTCAAAACCTTAAGCAATAGAGRRRTTCTTG		3240
Db	3181	TCCAGGCTATAGTGGCTTATCTTCATCTCAAAACCTTAAGCAATAGAGRRRTTCTTG		3240
Qy	3241	GCTTAATAGGTTTCTGCCAAMAATGANTCCCGAGGTTGGCRBAATAGCCAGATCA		3300
Db	3241	GCTTAATAGGTTTCTGCCAAMAATGANTCCCGAGGTTGGCRBAATAGCCAGATCA		3300
Qy	3301	WATCACTAATTAAGGAACTCAGAAAAGCAATCCCATTTATTAAGTAGATVAMCTGAA		3360
Db	3301	WATCACTAATTAAGGAACTCAGAAAAGCAATCCCATTTATTAAGTAGATVAMCTGAA		3360

```
Db 3301 WATACASTAATTAAAGAACTCAGAAAGCAATACCATTATKTAAGATGAYAMCTGA 3360
Qy 3361 GYMBAAGTGGCTTTCCAGGCCCTTAAGAGGCTTTAAACCAAGYCCAGATGTTAAAT 3420
Db 3361 GYMBAAGTGGCTTTCCAGGCCCTTAAGAGGCTTTAAACCAAGYCCAGATGTTAAAT 3420
Qy 3421 TGCCAAACGGGCAAGACTTTTSTYATAYRTCAAGAAAAAACAABAAVAGCTCTRGGA 3480
Db 3421 TGCCAAACGGGCAAGACTTTTSTYATAYRTCAAGAAAAAACAABAAVAGCTCTRGGA 3480
Qy 3481 GTCCCTTACACAGRTCCBAGGGAAGAGCTTGCAACCTRTGSGRYACCTAATAGGAAYT 3540
Db 3481 GTCCCTTACACAGRTCCBAGGGAAGAGCTTGCAACCTRTGSGRYACCTAATAGGAAYT 3540
Qy 3541 GATGATGCGCAAGGGGTTGRCYTCATGTTTAYGGGAGTGGGCGAGTGAAGTYTA 3600
Db 3541 GATGATGCGCAAGGGGTTGRCYTCATGTTTAYGGGAGTGGGCGAGTGAAGTYTA 3600
Qy 3601 GATCTGAAGCAAGTTAAATATATACAGGGAAGATCTTACTGTGTGACATCTCATGAK 3660
Db 3601 GATCTGAAGCAAGTTAAATATATACAGGGAAGATCTTACTGTGTGACATCTCATGAK 3660
Qy 3661 GTGAAYGCAATCTCACTGCTAAAGAGACTTGTGCGTGTGACAACTGTTACTTAA 3720
Db 3661 GTGAAYGCAATCTCACTGCTAAAGAGACTTGTGCGTGTGACAACTGTTACTTAA 3720
Qy 3721 TRTCAGGCTCTAATTAATCTTGAARGGCCAGTGTGCGTGTGACATCTTGAAC 3780
Db 3721 TRTCAGGCTCTAATTAATCTTGAARGGCCAGTGTGCGTGTGACATCTTGAAC 3780
Qy 3781 CCAAGYCNATTTCTTCCAGACATGAAAGAAAGATABAAYTAACTGTCAACATATAT 3840
Db 3781 CCAAGYCNATTTCTTCCAGACATGAAAGAAAGATABAAYTAACTGTCAACATATAT 3840
Qy 3841 TCTCAAACTATGCGCATCGAGGGGACTTGTAGARGTTCCTTGAATCTGATCCGACTT 3900
Db 3841 TCTCAAACTATGCGCATCGAGGGGACTTGTAGARGTTCCTTGAATCTGATCCGACTT 3900
Qy 3901 CAACCTGTATACGTATGGAAGTTCTTGTAGAAAAAGACTTGAAGYGGGGTATGC 3960
Db 3901 CAACCTGTATACGTATGGAAGTTCTTGTAGAAAAAGACTTGAAGYGGGGTATGC 3960
Qy 3961 AGTGCTCATGATATGGAATAYTTGAAGTATCCCTCATCTCAGAGAACTAGTCTTA 4020
Db 3961 AGTGCTCATGATATGGAATAYTTGAAGTATCCCTCATCTCAGAGAACTAGTCTTA 4020
Qy 4021 GCTRGCAAGACTAAATAGCCYTCATYKGGGCACTAGAAATTAGAGAAABAAAGGTAAA 4080
Db 4021 GCTRGCAAGACTAAATAGCCYTCATYKGGGCACTAGAAATTAGAGAAABAAAGGTAAA 4080
Qy 4081 TATATATACAGACTCTPAPTATGCTTACCTAGTCTCATGCGCATGCGCATGAGCAATATG 4140
Db 4081 TATATATACAGACTCTPAPTATGCTTACCTAGTCTCATGCGCATGCGCATGAGCAATATG 4140
Qy 4141 AGAAAGGAATTCCTTAATCTTGYAGAGAACCTATCAMAATCAGAGAAACCATATAGAR 4200
Db 4141 AGAAAGGAATTCCTTAATCTTGYAGAGAACCTATCAMAATCAGAGAAACCATATAGAR 4200
Qy 4201 ATTAATTAATGSGCWGTACAGAAACCTAABAGAGTGGMACTTCACTGCGGGGTCTCA 4260
Db 4201 ATTAATTAATGSGCWGTACAGAAACCTAABAGAGTGGMACTTCACTGCGGGGTCTCA 4260
Qy 4261 NAAAGGAAGRAAAGGAATTAAGRAAYTGCCAGCAKATTTTGAACMAAAGAGC 4320
Db 4261 NAAAGGAAGRAAAGGAATTAAGRAAYTGCCAGCAKATTTTGAACMAAAGAGC 4320
Qy 4321 TGCAAGGCAAGACCTTCATTAAGAAATGCTTATTAACCTTCCCTAGTATAGGTAATCC 4380
Db 4321 TGCAAGGCAAGACCTTCATTAAGAAATGCTTATTAACCTTCCCTAGTATAGGTAATCC 4380
Qy 4381 TTCCCGGGAACCAAGCCCAATGCTCAAGAGAGAAACGAATGGGGAACTTCAAGAG 4440
Db 4381 TTCCCGGGAACCAAGCCCAATGCTCAAGAGAGAAACGAATGGGGAACTTCAAGAG 4440
Qy 4441 CAGTTTCTCCCTCGGGAGCGTTAGCCATGGAAGAGGAAATATCTTTGCTGCAAC 4500
Db 4441 CAGTTTCTCCCTCGGGAGCGTTAGCCATGGAAGAGGAAATATCTTTGCTGCAAC 4500
Qy 4501 TATCAATGGAATTTACTTAAACCTTCAATCAAACTTTCACTTAGGCATGATAGAC 4560
Db 4501 TATCAATGGAATTTACTTAAACCTTCAATCAAACTTTCACTTAGGCATGATAGAC 4560
Qy 4561 CCATCAATGCGCAATATATTTACTGGAACGAGCTTTCAAACTATCAAGCAT 4620
Db 4561 CCATCAATGCGCAATATATTTACTGGAACGAGCTTTCAAACTATCAAGCAT 4620
Qy 4621 AKTCAGGGCTGTGAATGTGCGCABAATAATCCCTGCTAATGCGCAAGCTCTTC 4680
Db 4621 AKTCAGGGCTGTGAATGTGCGCABAATAATCCCTGCTAATGCGCAAGCTCTTC 4680
Qy 4681 AGABAAACAABAAACAGGCCATTACCTGABARAACCTGCAACTGATTTTACCAAG 4740
Db 4681 AGABAAACAABAAACAGGCCATTACCTGABARAACCTGCAACTGATTTTACCAAG 4740
Qy 4741 CCCAACTCAGGGAATTTCAATCTAATCTAATGCTGCTGCTTCACTTCAAGGGTGGCA 4800
Db 4741 CCCAACTCAGGGAATTTCAATCTAATCTAATGCTGCTGCTTCACTTCAAGGGTGGCA 4800
Qy 4801 RAGGCTTCCCCTGTAGGACAGAAAAAGCCCAAGAGTAAATTAAGGCACTATGTA 4860
Db 4801 RAGGCTTCCCCTGTAGGACAGAAAAAGCCCAAGAGTAAATTAAGGCACTATGTA 4860
Qy 4861 ATATATCCAGATTGCACTTCCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAG 4920
Db 4861 ATATATCCAGATTGCACTTCCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAG 4920
Qy 4921 GCCAAGTAAACCAAGGAGATCCAGGGCTTGTAGTATAGATATCACTTACCTGAGCC 4980
Db 4921 GCCAAGTAAACCAAGGAGATCCAGGGCTTGTAGTATAGATATCACTTACCTGAGCC 4980
Qy 4981 TGAAGGCAACAGTCTCAAGGGAAGTCAAGAAATGATGAAYACTCAAGGACATCTA 5040
Db 4981 TGAAGGCAACAGTCTCAAGGGAAGTCAAGAAATGATGAAYACTCAAGGACATCTA 5040
Qy 5041 AAAAGCAAAACCAAGGAAACCACTCAATGCGCTGTGTGCTTATAGCTTAA 5100
Db 5041 AAAAGCAAAACCAAGGAAACCACTCAATGCGCTGTGTGCTTATAGCTTAA 5100
Qy 5101 AGAATCTGCAATTTCCCAABAAAGAGACTTAAGCCATACGAATAGCTGTATGAGG 5160
Db 5101 AGAATCTGCAATTTCCCAABAAAGAGACTTAAGCCATACGAATAGCTGTATGAGG 5160
Qy 5161 CCCTCATTAACCAATGACCTTGTGTGACCAAGACAGCACTTAGTTCAGACATCA 5220
Db 5161 CCCTCATTAACCAATGACCTTGTGTGACCAAGACAGCACTTAGTTCAGACATCA 5220
Qy 5221 CCTCTTAGCCAAATATCAAAAGTTCTTAAACATTTAAGAGAACTATCCCTGAGAG 5280
Db 5221 CCTCTTAGCCAAATATCAAAAGTTCTTAAACATTTAAGAGAACTATCCCTGAGAG 5280
Qy 5281 AGGGAAGAAACATATTCACCCMWTGACATGATTAAGCAAGTCCCTCYCTAAT 5340
Db 5281 AGGGAAGAAACATATTCACCCMWTGACATGATTAAGCAAGTCCCTCYCTAAT 5340
Qy 5341 CCCCATCCCTAGATACATCTGAGAGAGCCCTACCTCATTTTATATTAACCCCACTG 5400
Db 5341 CCCCATCCCTAGATACATCTGAGAGAGCCCTACCTCATTTTATATTAACCCCACTG 5400
Qy 5401 CGGTTAAAGTGGCTGAGTGTAGTCTTGAATACATCACTTGAATCAAACTCTGATAC 5460
Db 5401 CGGTTAAAGTGGCTGAGTGTAGTCTTGAATACATCACTTGAATCAAACTCTGATAC 5460
Qy 5461 TGCAAAAGAACCTGAAATTCAGAGAGCAAGCTAGTATCTGCTGAGAG 5520
Db 5461 TGCAAAAGAACCTGAAATTCAGAGAGCAAGCTAGTATCTGCTGAGAG 5520
```

OY	5521	ATTGCGCTGCTCTTCAAACAACACGAGAGAAAGTAATCTAAATCTAAATCCCC	5580
Db	5521	ATTGCGCTGCTCTTCAAACAACACGAGAGAAAGTAATCTAAATCTAAATCCCC	5580
OY	5581	ATGAGCCCTCCCTAATCATATTTTCTCKTASTGTSHTTAACCTSTTTCACTCACT	5640
Db	5581	ATGAGCCCTCCCTAATCATATTTTCTCKTASTGTSHTTTTAACTSTTTCACTCACT	5640
OY	5641	GCACCCCTCCATGCGGTGTATGACAGTAGTCCCTYACOMAGATTCTATGAGAG	5700
Db	5641	GCACCCCTCCATGCGGTGTATGACAGTAGTCCCTYACOMAGATTCTATGAGAG	5700
OY	5701	ATGCAAGCGTCCCGAAATATTGATGCCCCCATCTGATAGAGACTTTTSTAAGGAAACCCC	5760
Db	5701	ATGCAAGCGTCCCGAAATATTGATGCCCCCATCTGATAGAGACTTTTSTAAGGAAACCCC	5760
OY	5761	ACCTTCACMGCCACACCCATATGCCCCGCAACTGCATCACTCTGCACTCTTTGCATG	5820
Db	5761	ACCTTCACMGCCACACCCATATGCCCCGCAACTGCATCACTCTGCACTCTTTGCATG	5820
OY	5821	CATGCAATACATCATTTATTTGACAGGAAATAATGATTAACTTAATTCTCGAGAGACTT	5880
Db	5821	CATGCAATACATCATTTATTTGACAGGAAATAATGATTAACTTAATTCTCGAGAGACTT	5880
OY	5881	GGAGTCACGTCTGTGTGACTTACTTCAACCCAACTGGTATGTGATGGGGGTGAAGTT	5940
Db	5881	GGAGTCACGTCTGTGTGACTTACTTCAACCCAACTGGTATGTGATGGGGGTGAAGTT	5940
OY	5941	CAAGATCAGGCAAGAAACATGTAAAGATATCTCCCACTCACCSGGGTACAT	6000
Db	5941	CAAGATCAGGCAAGAAACATGTAAAGATATCTCCCACTCACCSGGGTACAT	6000
OY	6001	GGCACTCTAGGCCCTTACAAAGGACTAATCTCAAACTACATGAAACCCCTCCATAC	6060
Db	6001	GGCACTCTAGGCCCTTACAAAGGACTAATCTCAAACTACATGAAACCCCTCCATAC	6060
OY	6061	CATATCTGCTGTGTAAAGCTATTATTAACAACCCCTCACTGGGCTTCATGAGTCTGGCC	6120
Db	6061	CATATCTGCTGTGTAAAGCTATTATTAACAACCCCTCACTGGGCTTCATGAGTCTGGCC	6120
OY	6121	CAAAACCTTAATCTGTGTGATATGCTCCCTCGAACTTCAAGCCATATGTTGCATC	6180
Db	6121	CAAAACCTTAATCTGTGTGATATGCTCCCTCGAACTTCAAGCCATATGTTGCATC	6180
OY	6181	CCTGATCCGGAACAATGGAACAACCTCGACAGAAATAAACAACAACCTCCGTTTATGTA	6240
Db	6181	CCTGATCCGGAACAATGGAACAACCTCGACAGAAATAAACAACAACCTCCGTTTATGTA	6240
OY	6241	GGACCTCTGTTCATATGTGAAATATGCAATACCTCAAACTTCACTGTGTAAATTT	6300
Db	6241	GGACCTCTGTTCATATGTGAAATATGCAATACCTCAAACTTCACTGTGTAAATTT	6300
OY	6301	AGCAATACTACATACACAACCACTCCCATATGATCAGGTGGGTAACTCTCCACACA	6360
Db	6301	AGCAATACTACATACACAACCACTCCCATATGATCAGGTGGGTAACTCTCCACACA	6360
OY	6361	ATAGTCTGCTACCCCTCAGGAAATATTTTGTGTGTGACTCAAGCTTATGTGTTGG	6420
Db	6361	ATAGTCTGCTACCCCTCAGGAAATATTTTGTGTGTGACTCAAGCTTATGTGTTGG	6420
OY	6421	AATGGCTCTTCAGAAATCTATGTCTCTCTCATTTCTTATGTCGCCCCYATGRCATCTAC	6480
Db	6421	AATGGCTCTTCAGAAATCTATGTCTCTCTCATTTCTTATGTCGCCCCYATGRCATCTAC	6480
OY	6481	ACTGAACAAGATTATATACGTTATGTCAATCTAAGCCCCCGTAAACAAAGATACCCATT	6540
Db	6481	ACTGAACAAGATTATATACGTTATGTCAATCTAAGCCCCCGTAAACAAAGATACCCATT	6540
OY	6541	CTTCCCTTTTGTATGAGACAGGAGGCTAAGGATGACTAAGTATGGCACTTGGCGGTATC	6600
Db	6541	CTTCCCTTTTGTATGAGACAGGAGGCTAAGGATGACTAAGTATGGCACTTGGCGGTATC	6600
OY	6601	ACAACTCTACTCAGTCTACTACTACAACTATCTCAAGACTAAATGGGACATGGAACGG	6660

Db	6601	ACAACTCTACTAGTTCTTACTCAAACTTCTCAAGACTAATGGGACATGGAAACGG	6660
Qy	6661	GTGCGCAGCTCCCTGTGCACCTTGCAGATCAACTTAACTCCCTAGACAGATAGTCCTT	6720
Db	6661	GTGCGCAGCTCCCTGTGCACCTTGCAGAAATCACTTAACTCCCTAGACAGATAGTCCTT	6720
Qy	6721	CTAAATGGAAGACTTTAGACTTTCGCTTAAACCGGCTGAGAGAGGGGAAACCTGTTATTTTA	6780
Db	6721	CTAAATGGAAGAGCTTTAGACTTTCGCTTAAACCGGCTGAGAGAGGGGAAACCTGTTATTTTA	6780
Qy	6781	GGGGAGAAATGCTGTTATTTATGTTAAATCAATCCGGAATGCTCACTGAGAAAGTTTAAAG	6840
Db	6781	GGGGAGAAATGCTGTTATTTATGTTAAATCAATCCGGAATGCTCACTGAGAAAGTTTAAAG	6840
Qy	6841	ATTGSAAGATCGAATPACACGCTAAGACAGAGAGACTTCGAAACCTGACCTGAGGCTC	6900
Db	6841	ATTGSAAGATCGAATPACACGCTAAGACAGAGAGACTTCGAAACCTGAGCCTGAGGCTC	6900
Qy	6901	CTCAGCCRAATGATGCCCTGGAATTCCTCTTCTTAGAGACTTACAGACTATATATTTG	6960
Db	6901	CTCAGCCRAATGATAGCCCTGGAATTCCTCTTCTTAGAGACTTACAGACTATATATTTG	6960
Qy	6961	CTATCTCTCTTTGGAGCCCTGATCTTTTACCTCTGTTAACTTGTCTTCCAGATC	7020
Db	6961	CTACTCTCTTTGGAGCCCTGATCTTTTACCTCTTGTAACTTGTCTTCCAGATC	7020
Qy	7021	GAAGCTGTAAATCTACAAATGGAGCCCAAGATCGACAACTAAGATCTACCGCAGA	7080
Db	7021	GAAGCTGTAAATCTACAAATGGAGCCCAAGATCGACAACTAAGATCTACCGCAGA	7080
Qy	7081	CCCCTGAGCCGGCTGTAGCCCAAGATCTGATGTTATATGACATCAAGGCAACCCCTCT	7140
Db	7081	CCCCTGAGCCGGCTGTAGCCCAAGATCTGATGTTATATGACATCAAGGCAACCCCTCT	7140
Qy	7141	GAGGAATTCACGCTGACAACTCTTACAGGCCCAATCAGAGGAACAGTTAGAGC	7200
Db	7141	GAGGAATTCACGCTGACAACTCTTACAGGCCCAATCAGAGGAACAGTTAGAGC	7200
Qy	7201	GGTGCTGAGGCAACCTCCCAACAGCACTTAAAGTTTCTGTGAGATGGGGGACTGAGA	7260
Db	7201	GGTGCTGAGGCAACCTCCCAACAGCACTTAAAGTTTCTGTGAGATGGGGGACTGAGA	7260
Qy	7261	GACAGGACTAGCTGATTTCTTAGGCTGATTAAGATTCCTTAAAGCTTAAAGCTGAGT	7320
Db	7261	GACAGGACTAGCTGATTTCTTAGGCTGATTAAGATTCCTTAAAGCTTAAAGCTGAGT	7320
Qy	7321	ACCAATTCACCTTTAAACCGGGGCTTGAACCTTATACACCTGACCAATCAGAGAG	7380
Db	7321	ACCAATTCACCTTTAAACCGGGGCTTGAACCTTATATACACCTGACCAATCAGAGAG	7380
Qy	7381	CTCACTAAATGCTAATTAGGCAAGACAGAGGTAAAGAAATAGCAATCATATATATTC	7440
Db	7381	CTCACTAAATGCTAATTAGGCAAGACAGAGGTAAAGAAATAGCAATCATATATATTC	7440
Qy	7441	MTGAGAGCAGACGAGAGGAGCAATGATCGGATATATAACCAAGTTCGAGCCGGCA	7500
Db	7441	MTGAGAGCAGACGAGAGGAGCAATGATCGGATATATAACCAAGTTCGAGCCGGCA	7500
Qy	7501	CGGCAACCCCTTTGGGTCCCTCTCCCTTGTATGGAGCTGTGTTTCACTATTTTAC	7560
Db	7501	CGGCAACCCCTTTGGGTCCCTCTCCCTTGTATGGAGCTGTGTTTCACTATTTTAC	7560
Qy	7561	TCTATTAATCTTGACACTGCR 7582	
Db	7561	TCTATTAATCTTGACACTGCR 7582	
RESULT 4			
AX027480			
LOCUS	AX027480	7582 bp	DNA
DEFINITION	Sequence 30 from Patent FR2788784.		linear
ACCESSION	AX027480		

VERSION AX027480.1 GI:10188444  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 Mallet, F., Voisard, C. and Paranhos, B.G.  
AUTHORS Patent: FR 2788784-A 30 28-JUL-2000;  
JOURNAL BIO MERIEUX (FR)  
FEATURES Location/Qualifiers  
source 1..7582  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 98.8%; Score 7489.6; DB 6; Length 7582;  
Best Local Similarity 100.0%; Pired. No. 0;  
Matches 7582; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CAACAATCGGGATATTAACCCAGGCAATTCAGCTGGGAGACAGAGCCGCCCTTTGGGTCC 60  
DB 1 CAACAATCGGGATATTAACCCAGGCAATTCAGCTGGGAGACAGAGCCGCCCTTTGGGTCC 60  
QY 61 CTTCCTTTGTATGGAGAGCTGTTTCAATGCTATTTCACTCTAATTAATCTTGGCACTGCA 120  
DB 61 CTTCCTTTGTATGGAGAGCTGTTTCAATGCTATTTCACTCTAATTAATCTTGGCACTGCA 120  
QY 121 CTCTTGTGTCATGTTTCTTAAGGCTGAGCTGAGCTTTTGTCTCACGCTCCACACTGC 180  
DB 121 CTCTTGTGTCATGTTTCTTAAGGCTGAGCTGAGCTTTTGTCTCACGCTCCACACTGC 180  
QY 181 TGTTCGCCACACCGGACACCTGCGCTGACCTCCACCTCGGATCCCGAGAGGCTC 240  
DB 181 TGTTCGCCACACCGGACACCTGCGCTGACCTCCACCTCGGATCCCGAGAGGCTC 240  
QY 181 TGTTCGCCACACCGGACACCTGCGCTGACCTCCACCTCGGATCCCGAGAGGCTC 240  
DB 181 TGTTCGCCACACCGGACACCTGCGCTGACCTCCACCTCGGATCCCGAGAGGCTC 240  
QY 241 CGCTGCTCTCTGATCCAGGAGCGCCCAATTCGCGCTCCCAATTGGGCTAAAGGCTTC 300  
DB 241 CGCTGCTCTCTGATCCAGGAGCGCCCAATTCGCGCTCCCAATTGGGCTAAAGGCTTC 300  
QY 241 CGCTGCTCTCTGATCCAGGAGCGCCCAATTCGCGCTCCCAATTGGGCTAAAGGCTTC 300  
DB 241 CGCTGCTCTCTGATCCAGGAGCGCCCAATTCGCGCTCCCAATTGGGCTAAAGGCTTC 300  
QY 301 CATGTGCTGCAACGCTAAGTGGCTGGTGTCTTAATTGAGCTGAACACTAATCACT 360  
DB 301 CATGTGCTGCAACGCTAAGTGGCTGGTGTCTTAATTGAGCTGAACACTAATCACT 360  
QY 361 GGGTTCATAGCTCTCTCTGTGACCCAGGCTTCTAATAAATACTAATACTTAACACA 420  
DB 361 GGGTTCATAGCTCTCTCTGTGACCCAGGCTTCTAATAAATACTAATACTTAACACA 420  
QY 421 TGGCCCAAGATTCATTCCTTGGAAATCGTAGGSCAAAGAACTCCAGGTCAAGAAATAC 480  
DB 421 TGGCCCAAGATTCATTCCTTGGAAATCGTAGGSCAAAGAACTCCAGGTCAAGAAATAC 480  
QY 481 GARGCTTGGCAACATCTTGGAAAGGCTCTACGCTCTTGAAGTGTTCACACCATC 540  
DB 481 GARGCTTGGCAACATCTTGGAAAGGCTCTACGCTCTTGAAGTGTTCACACCATC 540  
QY 541 TTGGGAGCTCTGTGACCAAGAGACCCCGGTGACATTTTGGCAACMSAAGGACATCC 600  
DB 541 TTGGGAGCTCTGTGACCAAGAGACCCCGGTGACATTTTGGCAACMSAAGGACATCC 600  
QY 601 MAAGTATGGGAAACGTTCCCGGCAAGACAAAGAGCCCTTAAGACGTAATCTTGABAAAT 660  
DB 601 MAAGTATGGGAAACGTTCCCGGCAAGACAAAGAGCCCTTAAGACGTAATCTTGABAAAT 660  
QY 661 TGGGAAATTTGACCTTCAGACACTTAAGAAAGAAAGACTTAATTTCTTCAAGTCC 720  
DB 661 TGGGAAATTTGACCTTCAGACACTTAAGAAAGAAAGACTTAATTTCTTCAAGTCC 720  
QY 721 GCTTGGCACTCTCGAGAGGAAATTAATACTTAACACATCTTACAGTGAATCTTTTG 780  
DB 721 GCTTGGCACTCTCGAGAGGAAATTAATACTTAACACATCTTACAGTGAATCTTTTG 780

QY 781 TAGAAAAGGCAATGAGTAGGCAATGACCAATGACCAATCTTCTTCAATTAAGACAA 840  
DB 781 TAGAAAAGGCAATGAGTAGGCAATGACCAATGACCAATCTTCTTCAATTAAGACAA 840  
QY 841 CTCACAATTAATGTAAGAAAGTGTATTTATGCTTACAGAAAGCTTCAGAGTCACTCC 900  
DB 841 CTCACAATTAATGTAAGAAAGTGTATTTATGCTTACAGAAAGCTTCAGAGTCACTCC 900  
QY 901 CTATCCAGAGATCCCGACCTCTTCCCAATTAAGAAAGGCTTCAAGGCTTCAAGGCT 960  
DB 901 CTATCCAGAGATCCCGACCTCTTCCCAATTAAGAAAGGCTTCAAGGCTTCAAGGCT 960  
QY 961 TCCAAAAGGAGATAGCAAAAAGGCTTAAACAGTAAACCAAGAGTGCATATTTCCCAAT 1020  
DB 961 TCCAAAAGGAGATAGCAAAAAGGCTTAAACAGTAAACCAAGAGTGCATATTTCCCAAT 1020  
QY 1021 TATGACCCCTCCCAAGCAGTGGAGAGAAATTCGGCCAGCCAGAGTGCATGTCT 1080  
DB 1021 TATGACCCCTCCCAAGCAGTGGAGAGAAATTCGGCCAGCCAGAGTGCATGTCT 1080  
QY 1081 TTTTCTCCAGACTTAAAGCAATTAAGAAAGCTTGAATTAATTTCTGATGATGAGAT 1140  
DB 1081 TTTTCTCCAGACTTAAAGCAATTAAGAAAGCTTGAATTAATTTCTGATGATGAGAT 1140  
QY 1141 GATGCTAATATGTTGTTTACAGGCTTAAAGCAATTTCTGATGATGAGAT 1200  
DB 1141 GATGCTAATATGTTGTTTACAGGCTTAAAGCAATTTCTGATGATGAGAT 1200  
QY 1201 ATATATGCTACCTGCTAATACAGACTAACCCTTAAGAGAGTGCATTAATCTC 1260  
DB 1201 ATATATGCTACCTGCTAATACAGACTAACCCTTAAGAGAGTGCATTAATCTC 1260  
QY 1261 AGCTGAGGTTTGGGATCTGCTGATCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1320  
DB 1261 AGCTGAGGTTTGGGATCTGCTGATCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1320  
QY 1321 GAAGGAAAGAAATGATTCCTCCACAGGCGAGGAGGCTTCCAGTCTASACCTCATTTG 1380  
DB 1321 GAAGGAAAGAAATGATTCCTCCACAGGCGAGGAGGAGGCTTCCAGTCTASACCTCATTTG 1380  
QY 1381 GGGACACAGAAATCACTAATCAGTGAATGAGTGTGCTGAGACATTTGCTAATCTGTGTC 1440  
DB 1381 GGGACACAGAAATCACTAATCAGTGAATGAGTGTGCTGAGACATTTGCTAATCTGTGTC 1440  
QY 1441 TASAAGATTAAGAAATCTAAGAAAGAAATCTAAGAAAGTCAATGATGATGATGATGAT 1500  
DB 1441 TASAAGATTAAGAAATCTAAGAAAGAAATCTAAGAAAGTCAATGATGATGATGATGAT 1500  
QY 1501 ACAAGGAGGAGGAGAAATCTTACTGCTTCTGAGAGACTAAGGAGGCTTGAAG 1560  
DB 1501 ACAAGGAGGAGGAGAAATCTTACTGCTTCTGAGAGACTAAGGAGGCTTGAAG 1560  
QY 1561 GAAAGCTGCTCTGTCTCACTGACTCTTCTGAGGCACTAATCTTAAAGCTTAAAGT 1620  
DB 1561 GAAAGCTGCTCTGTCTCACTGACTCTTCTGAGGCACTAATCTTAAAGCTTAAAGT 1620  
QY 1621 TATCACTGAGTCACTGAGCAATTAAGAAAGCTTCAAGGCTTCAAGGCTTCAAGGCT 1680  
DB 1621 TATCACTGAGTCACTGAGCAATTAAGAAAGCTTCAAGGCTTCAAGGCTTCAAGGCT 1680  
QY 1681 CAAAATTAAGAAACCTTATGAACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCT 1740  
DB 1681 CAAAATTAAGAAACCTTATGAACTTGAACCTTGAACCTTGAACCTTGAACCTTGAACCT 1740  
QY 1741 GAGCAGGCGGAAACAGGCAACGGAATTAAGAAAGGCTTGAACCTTGAACCTTGAACCT 1800  
DB 1741 GAGCAGGCGGAAACAGGCAACGGAATTAAGAAAGGCTTGAACCTTGAACCTTGAACCT 1800  
QY 1801 CAGGCAAGTGAATTTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1860  
DB 1801 CAGGCAAGTGAATTTGAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAG 1860  
QY 1861 TAGGCTTGTCTTCAAGTGGCTTCAAGGACCTTAAAGAAAGTGTCAAGTGAAG 1920

DB 1861 TAGGGCTTGGCTTCAGAGCGGCTTACAGAGCACCTTAAAAAGATTGTCCAAGTGAAG 1920  
QY 1921 TAAGCGGCCCTTGGTCATAGCCCTTATTTCAGAGGAATCACTGGAAGGCCCACTGCCC 1980  
DB 1921 TAAGCGGCCCTTGGTCATAGCCCTTATTTCAGAGGAATCACTGGAAGGCCCACTGCCC 1980  
QY 1981 CAGGGGCAAAAGTCTTTTGAATGACAGAGCCACTAACCAAGATGATCCAGACAGAGACTG 2040  
DB 1981 CAGGGGCAAAAGTCTTTTGAATGACAGAGCCACTAACCAAGATGATCCAGACAGAGACTG 2040  
QY 2041 AGGGTGGCTGGGGGCAAGCGCATCCCATGCGATCAACCTTCAAGAGCCCTGGGATGCTT 2100  
DB 2041 AGGGTGGCTGGGGGCAAGCGCATCCCATGCGATCAACCTTCAAGAGCCCTGGGATGCTT 2100  
QY 2101 GACCAATTGAGGGGCAAGCGCATCCCATGCGATCAACCTTCAAGAGCCCTGGGATGCTT 2160  
DB 2101 GACCAATTGAGGGGCAAGCGCATCCCATGCGATCAACCTTCAAGAGCCCTGGGATGCTT 2160  
QY 2161 CTTCCTGTCGGGCAACACTGTCCTTCAGATCTGTCACTATTCTGAGGGGGTCNTAAGACG 2220  
DB 2161 CTTCCTGTCGGGCAACACTGTCCTTCAGATCTGTCACTATTCTGAGGGGGTCNTAAGACG 2220  
QY 2221 GGGAGTCACTAGATTAATCTTCTTCCAGCGCACTAAGTTGAACCTGGGGAGCTTATTTCTT 2280  
DB 2221 GGGAGTCACTAGATTAATCTTCTTCCAGCGCACTAAGTTGAACCTGGGGAGCTTATTTCTT 2280  
QY 2281 TCACATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTATGGGAGAGACATTCTA 2340  
DB 2281 TCACATGCTTTTCTAATTATGCTTGAAGCCCACTACCTTGTATGGGAGAGACATTCTA 2340  
QY 2341 GCAAAAGCAGGGGCACTTAATACCTTGAACATAGAGAGAGAGACACCCGTTGTGTTGCC 2400  
DB 2341 GCAAAAGCAGGGGCACTTAATACCTTGAACATAGAGAGAGAGACACCCGTTGTGTTGCC 2400  
QY 2401 CCTGCTTGAAGAGGAATTAATCTGTAAGTGTGGGCAACAGAGAGACATATGAGCAGAC 2460  
DB 2401 CCTGCTTGAAGAGGAATTAATCTGTAAGTGTGGGCAACAGAGAGACATATGAGCAGAC 2460  
QY 2461 CAAGAATGCCCCGCTCTGTTCAAGTTAACTAAGATTCCACTTCTTCCCTACCAAA 2520  
DB 2461 CAAGAATGCCCCGCTCTGTTCAAGTTAACTAAGATTCCACTTCTTCCCTACCAAA 2520  
QY 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAAGATTCCAAAAGATTGTAAGGACTTAA 2580  
DB 2521 GGCAGTACCCCTCAGACCCCAAGGCCCAAGATTCCAAAAGATTGTAAGGACTTAA 2580  
QY 2581 AAGGCCAAGGTTAGTAAACATGCAATTAATCCCTGCAATTAATCCGTAAGTGAAG 2640  
DB 2581 AAGGCCAAGGTTAGTAAACATGCAATTAATCCCTGCAATTAATCCGTAAGTGAAG 2640  
QY 2641 GAGGCAAGAAACCAAGTGAAGTGAAGGTTAGTGAAGATCTCAGAGATTATCAATGG 2700  
DB 2641 GAGGCAAGAAACCAAGTGAAGTGAAGGTTAGTGAAGATCTCAGAGATTATCAATGG 2700  
QY 2701 AGGCGGTTGTCCTTTATATCCAGCTGTAACCTAGCCCTTATACCTGTGTTCCCAAATAC 2760  
DB 2701 AGGCGGTTGTCCTTTATATCCAGCTGTAACCTAGCCCTTATACCTGTGTTCCCAAATAC 2760  
QY 2761 CAGAGGAAGCAGATGTTTACATCTGAGACCTTGAAGATGCTTCTTCTGATCCTG 2820  
DB 2761 CAGAGGAAGCAGATGTTTACATCTGAGACCTTGAAGATGCTTCTTCTGATCCTG 2820  
QY 2821 TACATCTGATCTCAATTTCTTGTGCTTGAAGATTAATCAAAACCAACATCTCAAC 2880  
DB 2821 TACATCTGATCTCAATTTCTTGTGCTTGAAGATTAATCAAAACCAACATCTCAAC 2880  
QY 2881 TCACCTGAGCTRTTTTACCCCAAGGTTCAAGGATAGTCCCATCTATTTGGCCAGGAT 2940  
DB 2881 TCACCTGAGCTRTTTTACCCCAAGGTTCAAGGATAGTCCCATCTATTTGGCCAGGAT 2940  
QY 2941 TAGCCCAAGACTTGAATATTTATCTGGAACCTTGTCTTCTGTGATGAGATG 3000  
DB 2941 TAGCCCAAGACTTGAATATTTATCTGGAACCTTGTCTTCTGTGATGAGATG 3000

DB 2941 TAGCCCAAGACTTGAATATTTATCTGGAACCTTGTCTTCTGTGATGAGATG 3000  
QY 3001 ATTATCTTTAGCTGCGCTRTTCAAGAAACCTTGTGCAATCAAGCAACCAAGCTCTTAA 3060  
DB 3001 ATTATCTTTAGCTGCGCTRTTCAAGAAACCTTGTGCAATCAAGCAACCAAGCTCTTAA 3060  
QY 3061 ATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
DB 3061 ATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3120  
QY 3121 AGGTTAAATCTTAAAGCTTAAATTAATCAAAAGGCAACAGGCTTCAAGTGAAGATTA 3180  
DB 3121 AGGTTAAATCTTAAAGCTTAAATTAATCAAAAGGCAACAGGCTTCAAGTGAAGATTA 3180  
QY 3121 AGGTTAAATCTTAAAGCTTAAATTAATCAAAAGGCAACAGGCTTCAAGTGAAGATTA 3180  
DB 3121 AGGTTAAATCTTAAAGCTTAAATTAATCAAAAGGCAACAGGCTTCAAGTGAAGATTA 3180  
QY 3181 TCCAGCTTACTGCTGCTTATCCCATCYCAAAACCTTAAAGCACTAAGGRTTCTTGG 3240  
DB 3181 TCCAGCTTACTGCTGCTTATCCCATCYCAAAACCTTAAAGCACTAAGGRTTCTTGG 3240  
QY 3241 GCTTAAYAGGTTCTGCGCAAAATGATTCGCCAGGTGTCRBAATAGCCAGGYCATTA 3300  
DB 3241 GCTTAAYAGGTTCTGCGCAAAATGATTCGCCAGGTGTCRBAATAGCCAGGYCATTA 3300  
QY 3301 WATACASTAATTAAGAAATCTCAAGAAACCAATCCATTTATTAAGTGAATAMCTGA 3360  
DB 3301 WATACASTAATTAAGAAATCTCAAGAAACCAATCCATTTATTAAGTGAATAMCTGA 3360  
QY 3301 WATACASTAATTAAGAAATCTCAAGAAACCAATCCATTTATTAAGTGAATAMCTGA 3360  
DB 3301 WATACASTAATTAAGAAATCTCAAGAAACCAATCCATTTATTAAGTGAATAMCTGA 3360  
QY 3361 GYRBAAGTGGCTTCCAGGCCCCCTTAAAGAGGCTTAAACCCAGYCCAGTGTAAAGT 3420  
DB 3361 GYRBAAGTGGCTTCCAGGCCCCCTTAAAGAGGCTTAAACCCAGYCCAGTGTAAAGT 3420  
QY 3421 TGCCAAACRGGGCAAGCTTTTSTTATATCAAGAAAAAAGABAAVAGCTCTGGA 3480  
DB 3421 TGCCAAACRGGGCAAGCTTTTSTTATATCAAGAAAAAAGABAAVAGCTCTGGA 3480  
QY 3481 GTCCTTACACAGTCCRAAGGAGAGAGCTTGAACCTRTGCGRACTGATGAAGAAAT 3540  
DB 3481 GTCCTTACACAGTCCRAAGGAGAGAGCTTGAACCTRTGCGRACTGATGAAGAAAT 3540  
QY 3541 GATGTAGTGGCAAAAGGTTGTCATTTGTTTAYGGGTAGTGGGCAAGTGAAGTGA 3600  
DB 3541 GATGTAGTGGCAAAAGGTTGTCATTTGTTTAYGGGTAGTGGGCAAGTGAAGTGA 3600  
QY 3601 GATGTAGTGGCAAAAGGTTGTCATTTGTTTAYGGGTAGTGGGCAAGTGAAGTGA 3660  
DB 3601 GATGTAGTGGCAAAAGGTTGTCATTTGTTTAYGGGTAGTGGGCAAGTGAAGTGA 3660  
QY 3661 GTATCTGAAGCATTAATAATATATACAGGAGAGATCTTACTGTGAGACATCTCAAGAK 3720  
DB 3661 GTATCTGAAGCATTAATAATATATACAGGAGAGATCTTACTGTGAGACATCTCAAGAK 3720  
QY 3721 TRTCAAGCTCTATTAATCTGTAAGGCGAGTGTGCTGCAAGCACTGTTAACTTAA 3780  
DB 3721 TRTCAAGCTCTATTAATCTGTAAGGCGAGTGTGCTGCAAGCACTGTTAACTTAA 3780  
QY 3781 CCAGYCAATTTCTTCCAGCAAAATGAAGAAAGATTAATATCTGTCAACATTAAT 3840  
DB 3781 CCAGYCAATTTCTTCCAGCAAAATGAAGAAAGATTAATATCTGTCAACATTAAT 3840  
QY 3841 TCTCAAACTATGCTCACTGAGGGGACCTTATGAAGTTCCTTGAAGTTCCTTGAAGT 3900  
DB 3841 TCTCAAACTATGCTCACTGAGGGGACCTTATGAAGTTCCTTGAAGTTCCTTGAAGT 3900  
QY 3901 CAACTTGTATATGATGAAGGTTCTTGTGTAAGAAAAAGACTTGAAGAAAGTGGGATAC 3960  
DB 3901 CAACTTGTATATGATGAAGGTTCTTGTGTAAGAAAAAGACTTGAAGAAAGTGGGATAC 3960  
QY 3961 AGTGTAGTATATGATGAAGTTCCTTGTGTAAGAAAAAGACTTGAAGAAAGTGGGATAC 4020  
DB 3961 AGTGTAGTATATGATGAAGTTCCTTGTGTAAGAAAAAGACTTGAAGAAAGTGGGATAC 4020  
QY 4021 GCTRGCAAGAACTAATAGCTTCAATGAGGCACTGAGATTAAGGAAAGGAAAGGATTA 4080  
DB 4021 GCTRGCAAGAACTAATAGCTTCAATGAGGCACTGAGATTAAGGAAAGGAAAGGATTA 4080  
QY 4081 GCTRGCAAGAACTAATAGCTTCAATGAGGCACTGAGATTAAGGAAAGGAAAGGATTA 4080  
DB 4081 GCTRGCAAGAACTAATAGCTTCAATGAGGCACTGAGATTAAGGAAAGGAAAGGATTA 4080



QY 4081 TATATATACAGACTCTBARTATGCTVACCTAGTCNTCCATGCCCATGMRCAATATGTSAR 4140  
Db 4081 TATATATACAGACTCTBARTATGCTVACCTAGTCNTCCATGCCCATGMRCAATATGTSAR 4140  
QY 4141 AGAAAGGAAATTCCTAACTTCYGAGRACAACCTATCAMAATCAGAGAGCCATTAGAR 4200  
Db 4141 AGAAAGGAAATTCCTAACTTCYGAGRACAACCTATCAMAATCAGAGAGCCATTAGAR 4200  
QY 4201 ATTATATYGGGCGTACAGAAACCTABAGAGTGGMAGTCTTACACTGCGGGGTCACTCA 4260  
Db 4201 ATTATATYGGGCGTACAGAAACCTABAGAGTGGMAGTCTTACACTGCGGGGTCACTCA 4260  
QY 4261 NAAAGGAAAGRAAGGAAATASAGRAAYTGCCAGACAKATTTAGACMAAAAGAGC 4320  
Db 4261 NAAAGGAAAGRAAGGAAATASAGRAAYTGCCAGACAKATTTAGACMAAAAGAGC 4320  
QY 4321 TGCAGAGCAGAGCCCTCCATTAGAAATGCTTATTTAACTTCCCTAGTATAGGTAATCC 4380  
Db 4321 TGCAGAGCAGAGCCCTCCATTAGAAATGCTTATTTAACTTCCCTAGTATAGGTAATCC 4380  
QY 4381 CTTCCGGGAAACCAAGCCCAAGTACTCAGACAGAGAAACAGATGGGAACTCAGAGG 4440  
Db 4381 CTTCCGGGAAACCAAGCCCAAGTACTCAGACAGAGAAACAGATGGGAACTCAGAGG 4440  
QY 4441 CAGTTTCTCCCTCGGAGCGTTAGGCACTGAAGAAAGGAAATACTTTTGCTGCTGAC 4500  
Db 4441 CAGTTTCTCCCTCGGAGCGTTAGGCACTGAAGAAAGGAAATACTTTTGCTGCTGAC 4500  
QY 4501 TATCCAAATGAAATTAATTAAACCTTCATCAAACTTTCACTTAGGCAATGATAGAC 4560  
Db 4501 TATCCAAATGAAATTAATTAAACCTTCATCAAACTTTCACTTAGGCAATGATAGAC 4560  
QY 4561 CCATCBAATGGCCAAATCAATTAATTACTGACAGAGCCTTTCAAAACTATCAAGABAT 4620  
Db 4561 CCATCBAATGGCCAAATCAATTAATTACTGACAGAGCCTTTCAAAACTATCAAGABAT 4620  
QY 4621 ACTCAGGCGCTGTGAATGTGACABARAAATTAATCCCTGCTATAGGCCAAGCTCTTC 4680  
Db 4621 ACTCAGGCGCTGTGAATGTGACABARAAATTAATCCCTGCTATAGGCCAAGCTCTTC 4680  
QY 4681 AGGARAACAARAACAAGGCCATTACCTGABARABATGSCAACTGATTTTACCCAAG 4740  
Db 4681 AGGARAACAARAACAAGGCCATTACCTGABARABATGSCAACTGATTTTACCCAAG 4740  
QY 4741 CCCAAACCTCAGGAGTTTCACTATCTAGTCTGGGTAAATACCTTCAAGGGTGGGCA 4800  
Db 4741 CCCAAACCTCAGGAGTTTCACTATCTAGTCTGGGTAAATACCTTCAAGGGTGGGCA 4800  
QY 4801 RAGGCTTCCCTGTAGGAAGAAAGGCCAAGAGGTAAATAAGGCACTAGTCAATGAA 4860  
Db 4801 RAGGCTTCCCTGTAGGAAGAAAGGCCAAGAGGTAAATAAGGCACTAGTCAATGAA 4860  
QY 4861 ATATATTCAGATTCGGAATTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTCCAG 4920  
Db 4861 ATATATTCAGATTCGGAATTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTCCAG 4920  
QY 4921 GCCACAGTAAACCAAGGAGTATCCCAAGGCTTATAGTATACATATCACTTACCTGGCC 4980  
Db 4921 GCCACAGTAAACCAAGGAGTATCCCAAGGCTTATAGTATACATATCACTTACCTGGCC 4980  
QY 4981 TGAAGGCCACAGTCTCAGGGAAGTGCAGAAATGAATGAAYACTCAAGAGCAATCTA 5040  
Db 4981 TGAAGGCCACAGTCTCAGGGAAGTGCAGAAATGAATGAAYACTCAAGAGCAATCTA 5040  
QY 5041 AAAAGCAAAACCAAGAAACCACTCACAATGAGCTGTCTGTGCTATAGCTTAAAA 5100  
Db 5041 AAAAGCAAAACCAAGAAACCACTCACAATGAGCTGTCTGTGCTATAGCTTAAAA 5100  
QY 5101 AGAATCTGCAATTTTCCCAAAAAGAGACTTACGCAATACGAATGCTGTATGGAAG 5160  
Db 5101 AGAATCTGCAATTTTCCCAAAAAGAGACTTACGCAATACGAATGCTGTATGGAAG 5160

QY 5161 CCCTTCATACCAATGACCTGTGTGACCCCAAGACAGCCAATTAGTGCAGACATCA 5220  
Db 5161 CCCTTCATACCAATGACCTGTGTGACCCCAAGACAGCCAATTAGTGCAGACATCA 5220  
QY 5221 CCTCTTACCAATATCAACAAGTTCTTAAACATTACAGAGACCTATCCTGAGAG 5280  
Db 5221 CCTCTTACCAATATCAACAAGTTCTTAAACATTACAGAGACCTATCCTGAGAG 5280  
QY 5281 AGGAAAGAAACATATCCACCCMGTGACATGGTATTAAGCAAGTCCCTTCCTATAT 5340  
Db 5281 AGGAAAGAAACATATCCACCCMGTGACATGGTATTAAGCAAGTCCCTTCCTATAT 5340  
QY 5341 CCCCATCCCTAGATACATCTGGGAGAGACCTTACCAGTCAATTTATYTAACCCAATG 5400  
Db 5341 CCCCATCCCTAGATACATCTGGGAGAGACCTTACCAGTCAATTTATYTAACCCAATG 5400  
QY 5401 CGGTTAAAGTGGCTGAGATGAGTCTTGATATACATCACTTGAGTCAAAATCCTGATAC 5460  
Db 5401 CGGTTAAAGTGGCTGAGATGAGTCTTGATATACATCACTTGAGTCAAAATCCTGATAC 5460  
QY 5461 TGCAGAAAGAACTGAAATTCAGAGAGCAACGCTAGTATTCCTGTGAACCTCTAGAG 5520  
Db 5461 TGCAGAAAGAACTGAAATTCAGAGAGCAACGCTAGTATTCCTGTGAACCTCTAGAG 5520  
QY 5521 ATTTGCGCTGCTCTTCAACAACAACAAGAGAAAGTAACTTAAATCAATAATCCCC 5580  
Db 5521 ATTTGCGCTGCTCTTCAACAACAACAAGAGAAAGTAACTTAAATCAATAATCCCC 5580  
QY 5581 ATGGGCTCTCTTATCATATTTTCTCTATATGTTSTTTTACCTCTTCACTCTCACT 5640  
Db 5581 ATGGGCTCTCTTATCATATTTTCTCTATATGTTSTTTTACCTCTTCACTCTCACT 5640  
QY 5641 GCACCCCTCCATGCGCTGTATGACAGTAAGTCCCTTACCMAGATTTCTATGAGA 5700  
Db 5641 GCACCCCTCCATGCGCTGTATGACAGTAAGTCCCTTACCMAGATTTCTATGAGA 5700  
QY 5701 ATGACGCGTCCGGAAATATGATGCGCCATGATAGAGTCTTSTAAGGAAACCCC 5760  
Db 5701 ATGACGCGTCCGGAAATATGATGCGCCATGATAGAGTCTTSTAAGGAAACCCC 5760  
QY 5761 ACCTTCACTGCCCAACCAATATGCCCGGCACTGCTATCACTCTGCGCACTCTTGATG 5820  
Db 5761 ACCTTCACTGCCCAACCAATATGCCCGGCACTGCTATCACTCTGCGCACTCTTGATG 5820  
QY 5821 CANTGAAATATCATTAATGACAGGAAATATGATTAATCTAGTGTCTGTGAGGACTT 5880  
Db 5821 CANTGAAATATCATTAATGACAGGAAATATGATTAATCTAGTGTCTGTGAGGACTT 5880  
QY 5881 GAGTCACTGTCTGTGACATTAATCTTCAACCAACTGTATGCTGATGGGGTGAAGTT 5940  
Db 5881 GAGTCACTGTCTGTGACATTAATCTTCAACCAACTGTATGCTGATGGGGTGAAGTT 5940  
QY 5941 CAAGATCAGGCAAGGAAACATGTAAGAAAGTATCTCCCACTCAGCCSGGTACAT 6000  
Db 5941 CAAGATCAGGCAAGGAAACATGTAAGAAAGTATCTCCCACTCAGCCSGGTACAT 6000  
QY 6001 GGACCTCAGGCGCTCAACAAGGACTGATCTCAAAACTACATGAAACCTCCGTAAC 6060  
Db 6001 GGACCTCAGGCGCTCAACAAGGACTGATCTCAAAACTACATGAAACCTCCGTAAC 6060  
QY 6061 CATATCGGCTGTGAAGCTATTTAATACACCTTCACTGGGCTCAGTGAAGTCTGCGCC 6120  
Db 6061 CATATCGGCTGTGAAGCTATTTAATACACCTTCACTGGGCTCAGTGAAGTCTGCGCC 6120  
QY 6121 CAATAACCTTACTTACTGTGATATGCTCTCCCTGGAATTTCAAGCAATATGTTCAATC 6180  
Db 6121 CAATAACCTTACTTACTGTGATATGCTCTCCCTGGAATTTCAAGCAATATGTTCAATC 6180  
QY 6181 CCTGTACCTGAAACATGGAACAATTCAGACAGAAATTAACCACTTCCGTTTATGTA 6240  
Db 6181 CCTGTACCTGAAACATGGAACAATTCAGACAGAAATTAACCACTTCCGTTTATGTA 6240  
QY 6241 GGAAGCTCTTGTTCATATSTGAAATTAACCACTACCTCAACCTCACTGTGTAAATTT 6300

Db 6241 GGAAGCTCTGTTTCCAAATGGAATTAACCACTCAACCTCAACCTCTGTGTAATTT 6300  
 QY 6301 ACCAATACATACATACACAACAACCACTCCCAATGATGAGGTGTAATCTCTCCACACA 6360  
 Db 6301 AGCAATACATACATACACAACAACCACTCCCAATGATGAGGTGTAATCTCTCCACACA 6360  
 QY 6361 ATAGTGTGCTTACCTCCAGGAATATTTTGTCTGTGTGTCTGTGTGTCTGTGTGTGT 6420  
 Db 6361 ATAGTGTGCTTACCTCCAGGAATATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6420  
 QY 6421 AATGCTCTTCAAGATCTATGTGTCTCTCAATCTTGTGTGTGTGTGTGTGTGTGTGTGT 6480  
 Db 6421 AATGCTCTTCAAGATCTATGTGTCTCTCAATCTTGTGTGTGTGTGTGTGTGTGTGTGT 6480  
 QY 6481 ACTGAACAAGATTATACAGTATGTATATCTAAGCCCGCAACAAGATGATCCATT 6540  
 Db 6481 ACTGAACAAGATTATACAGTATGTATATCTAAGCCCGCAACAAGATGATCCATT 6540  
 QY 6541 CTTCCTTTTGTATAG 6600  
 Db 6541 CTTCCTTTTGTATAG 6600  
 QY 6601 ACAACCTTACAGTGTCTTCTCAATCTAATCTCAAGAACTAATGGGAGATGAAACGG 6660  
 Db 6601 ACAACCTTACAGTGTCTTCTCAATCTAATCTCAAGAACTAATGGGAGATGAAACGG 6660  
 QY 6661 GTGCGCAGACTCCCTGTGATCTTGAAGATCAATTAATCTCCCTAGAGAGAGAGAGAG 6720  
 Db 6661 GTGCGCAGACTCCCTGTGATCTTGAAGATCAATTAATCTCCCTAGAGAGAGAGAGAG 6720  
 QY 6721 CCAATTCGAAGAGCTTGTAGACTGTCTAACCCTGAGAGAGAGAGAGAGAGAGAGAGAG 6780  
 Db 6721 CCAATTCGAAGAGCTTGTAGACTGTCTAACCCTGAGAGAGAGAGAGAGAGAGAGAGAG 6780  
 QY 6781 GGGGAAGAAAGCTGTAT 6840  
 Db 6781 GGGGAAGAAAGCTGTAT 6840  
 QY 6841 ATTCSAGATGAAATACAACTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6900  
 Db 6841 ATTCSAGATGAAATACAACTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6900  
 QY 6901 CTGAGCCGATGAG 6960  
 Db 6901 CTGAGCCGATGAG 6960  
 QY 6961 CTACTCTCTTTTGAAG 7020  
 Db 6961 CTACTCTCTTTTGAAG 7020  
 QY 7021 GAAGCTGTAAATCTAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7080  
 Db 7021 GAAGCTGTAAATCTAACAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7080  
 QY 7081 CCCCTGAG 7140  
 Db 7081 CCCCTGAG 7140  
 QY 7141 GAGGAATCTCAGCTGACAACTCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7200  
 Db 7141 GAGGAATCTCAGCTGACAACTCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7200  
 QY 7201 GGTGTGTGAG 7260  
 Db 7201 GGTGTGTGAG 7260  
 QY 7261 GACAG 7320  
 Db 7261 GACAG 7320  
 QY 7321 ACCAGATCCAGCTTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7380  
 Db 7321 ACCAGATCCAGCTTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7380

Db 7321 ACCAGATCCAGCTTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7380  
 QY 7381 CTCCTAAATAGCTTAATTAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7440  
 Db 7381 CTCCTAAATAGCTTAATTAAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7440  
 QY 7441 MTGAG 7500  
 Db 7441 MTGAG 7500  
 QY 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 7560  
 Db 7501 CGGCAACCCCTTTGGGTCCCTCCCTTTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 7560  
 QY 7561 TCTATTAATCTTGACAGCTGCR 7582  
 Db 7561 TCTATTAATCTTGACAGCTGCR 7582  
 RESULT 5  
 AY101582 10222 bp DNA linear PRI 11-FEB-2004  
 LOCUS Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVWE1 locus,  
 DEFINITION allele A, complete sequence.  
 ACCESSION AY101582  
 VERSION AY101582.1 GI:37544399  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 10222)  
 Mallet, P., Boulton, O., Prudhomme, S., Cheynet, V., Oriol, G.,  
 Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.  
 The endogenous retroviral locus ERVWE1 is a bona fide gene involved  
 in hominoid placental physiology  
 Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)  
 JOURNAL PUBMED 14757826  
 TITLE 2 (bases 1 to 10222)  
 Mallet, P., Boulton, O. and Oriol, G.  
 Direct Substition  
 Submitted (06-MAY-2002) Retrovirology Department, UMR 2142  
 CNRS-Biomerieux, Ecole Normale Supérieure de Lyon - 46 allée  
 d'Italie, Lyon 69634 cedex 07, France  
 FEATURES  
 source  
 1..10222  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /isolate="21"  
 /isolation\_source="PBMC of individual 21"  
 /db\_xref="taxon:9606"  
 /db\_xref="taxon:9606"  
 /chromosome="7"  
 /map="7q21-q22"  
 /sex="female"  
 /cell\_type="PBMC"  
 /note="isolated by PCR; ERVWE1 locus, allele A  
 endogenous\_virus="HERV-W"  
 1..780  
 join(248..811,2887..9813)  
 /note="putative mRNA transcrip 1; gag-pol pseudogene"  
 join(248..811,7585..9813)  
 /product="envelope glycoprotein"  
 /note="putative mRNA transcrip 2"  
 join(248..811,9248..9813)  
 /note="putative mRNA transcrip 3; no coding region"  
 7812..9428  
 /note="syncytin"  
 /codon\_start=1  
 /product="envelope glycoprotein"  
 /protein\_id="AA068161.1"  
 /db\_xref="GI:37544400"  
 /translation="MALPHYILFTVLALBSFTLTAPPCCMTSSSPYQEFRLMORP  
 GNIDAPSYSLSKGPTPTAHTMPNRCVSHATLCMHANHTHTGKMINPSCPGGLG

LTR  
ORIGIN

Query Match 85.0%; Score 6448; DB 9; Length 10222;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;

TVCMVYFTQIGMSDGGVDDOAREKHEKEVISQLTREVHGTSFHKGLDLSKHELTURT  
HRLVSLFPTTLTGLHEVSKONPTNCMI CLPLNRPVPSIPVEQNNFSTENITVS  
LVQPLVSNLEIHTISNLTCKVFNSTYITNSOCIRWTPPTQVCLPSGIFVQGSIA  
YCLNGSSSMCLSLFLVPMPTITBODLYSYISKPRNRVYIILPVPVCGYALG  
TIGGGITTSQFTYKJASOELNDEMERYASLVTLQDOLNSLA VYLQNRADULTAE  
RGQTCILFLEBECYVNOGSI VTEKVEKLRDRIOREBELRNTPGKLLSQWPMILP  
FGPLPAIILLEFGPCIFNLVNFVYSRIEAVYKQMEPMOSKTKIYRRPLDRPASP  
RSDVNDIKSTPEEISAAQPLLRPNAGSS"

9487..10222

606 GATGGGAAACGTTCCCGGAGCAAAAACGCCCTTAAGCGTATTCTGGAAATTTGGGA 665  
2886 GATGGGAAACGTTCCCGGAGCAAAAACGCCCTTAAGCGTATTCTGGAAATTTGGGA 2945  
666 MCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTAATTTCTTCGACGTCCGCTG 725  
2946 CCAATTTGACCCCTCAGACACTAAGAAAGAAACGACTTAATTTCTTCGACGTCCGCTG 3005  
726 GCACTCTGAGGAGAACTAATTAATTAACACCACTTAAGACTAGACCTTTTGTAGAA 785  
3006 GCACTCTGAGGAGAACTAATTAATTAACACCACTTAAGACTAGACCTTTTGTAGAA 3065  
786 AAGGAAATGAGAGTGAAGTCCATAGTACAACTTTCTTTCAATTAAGACAACTCAC 845  
3066 AAGGAAATGAGAGTGAAGTCCATAGTACAACTTTCTTTCAATTAAGACAACTCAC 3125  
846 AATTATGTAAAAAGTGTGATTTATGCCCTACAGAAAGCCCTTCAGAGTCACTCCCTATC 905  
3126 AATTATGTAAAAAGTGTGATTTATGCCCTACAGAAAGCCCTTCAGAGTCACTCCCTATC 3185  
906 CCAGCATCCCGCATCTCTTCCCAATTAAGAGACCCCTTCAACCAATGTCCTCA 965  
3186 CCAGCATCCCGCATCTCTTCCCAATTAAGAGACCCCTTCAACCAATGTCCTCA 3245  
966 AAGGATATGACAAAAGGTTAAACAGTAAACCAAGAGTCCCAATTTCCCAATTAAGA 1025  
3246 AAGGATATGACAAAAGGTTAAACAGTAAACCAAGAGTCCCAATTTCCCAATTAAGA 3305  
1026 CCCCTCCCAAGCAGTGGAGAGAAAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTY 1085  
3306 CCCCT-CCAAAGCAGTGGAGAGAAAGAAATTCGGCCCAAGCAGAGTGCATGTGCTTTTC 3364  
1086 TCTCCCAAGCTTAAGCAATTAATAACAGACTTAAGTAAATTTCTCAGATTAATCTGTATG 1145  
3365 TCTCCCAAGCTTAAGCAATTAATAACAGACTTAAGTAAATTTCTCAGATTAATCTGTATG 3424  
1146 CTATATGTGTGTTTAAAGAGTTAGAGCAATTTCTTGATCTGACATGAGAGATTAATA 1205  
3425 CTATATGTGTGTTTAAAGAGTTAGAGCAATTTCTTGATCTGACATGAGAGATTAATA-A 3483  
1206 TGTCACTGCTAATCAGACACTTAACCCCAATGAGAGAGTGCACCATTAATCTGACGCT 1265  
3484 TGTCACTGCTAATCAGACACTTAACCCCAATGAGAGAGTGCACCATTAATCTGACGCT 3543  
1266 GAGGTTTGGGAGATCTCTGTATCTCAGTCAAGTCAATGATAMGATTAACACAGAGG 1325  
3544 GAGGTTTGGGAGATCTCTGTATCTCAGTCAAGTCAATGATAMGATTAACAGAGG-AGG 3600  
1326 AAGGAAATGATTTCCCAAGCCAGCAGAGCAAGTTCCTCCAGTCTAASACCTCAATTGGGAGC 1385  
3601 AAGGAAATGATTTCCCAAGCCAGCAGAGCAAGTTCCTCCAGTCTAASACCTCAATTGGG-- 3657  
1386 ACAGAAATCATGTAACATGGAGATGTTGCTCAGACATTTGCTAACTGTGTGCTAASA 1445  
3658 ACAGAAATCATGAACT-AGAGATTTGCTCAGACATTTGCTAATCTGTGTGCTAASA 3716  
1446 GGAATTAAGGAAATCTAAGAAATCTAATTAATTAATGATGTCACATTAACACA 1505

3717 GGAATTAAGGAAATCTAAGAG-AGGCTATGATTAATCAATGATGTCACCATTAACACA 3775  
1506 GGGGAAAGGAAAGAAATTCCTAATGCTCTTCTGAGAGACTAAGGAGGATTAAGAGAGC 1565  
3776 -GGGAAAGGAAAGAAATTCCTAATGCTCTTCTGAGAGACTAAGGAGGATTAAGAGAGC 3834  
1566 GTGCTCTCTGCTCAGCTGACTCTTCTGAGAGGCACTAATCTTAAGCTTAAGTTATCA 1625  
3835 GTGCTCTCTGCTCAGCTGACTCTTCTGAGAGGCACTAATCTTAAGCTTAAGTTATCA 3894  
1626 CTGAGTCACTGAGACATTAAG-AAAAAATTAAGTCTGCGCTGAGGCGCGGAGCAAA 1684  
3895 CTGAGTCACTGAGACATTAAGAAAAAATTAAGTCTGCGCTGAGGCGCGGAGCAAA 3954  
1685 ACTTAAGAAACCTTAATGAACCTTGGGAACTGCGGTTTTTAATATAGATCAAGAGAGC 1744  
3955 ACTTAAGAAACCTTAATGAACCTTGGGAACTGCGGTTTTTAATATAGATCAAGAGAGC 4014  
1745 AGCGGAAACAGGACAAAGGAGTTAAAAAAGGCAAGCTTTAGTCAAGACCTCAGG 1804  
4015 AGCGGAAACAGGACAAAGGAGTTAAAAAAGGCAAGCTTTAGTCAAGACCTCAGG 4074  
1805 CAAGTGAACCTTGAAGGCTCTGGAAGAGGAAAGCTGGGCAATTTGAATGCTTAATAG 1864  
4075 CAAGTGAACCTTGAAGGCTCTGGAAGAGGAAAGCTGGGCAATTTGAATGCTTAATAG 4134  
1865 GCTTCTTCAAGTGGGCTCTAACAAGACACTTAATAAAGATTTGTCAGATGAAGTAA 1924  
4135 GCTTCTTCAAGTGGGCTCTAACAAGACACTTAATAAAGATTTGTCAGATGAAGTAA 4194  
1925 CGGCCCTTCTGTCATGCCCCCTTAATTAAGGAAATCACTGGAAGGCCCATGCCCCAG 1984  
4195 CGGCCCTTCTGTCATGCCCCCTTAATTAAGGAAATCACTGGAAGGCCCATGCCCCAG 4254  
1985 GGAACAAGGCTTTTGAATCAGAAAGCACTAACCAAGATGATCCACAGAGAGCTGAAGG 2044  
4255 GGAACAAGGCTTTTGAATCAGAAAGCACTAACCAAGATGATCCACAGAGAGCTGAAGG 4314  
2045 TGCCCTGGGCAAGCGCATCCCATGCACTGCAAGAGCCCTGGGATATGCTTGAAC 2104  
4315 TGCCCTGGGCAAGCGCATCCCATGCACTGCAAGAGCCCTGGGATATGCTTGAAC 4374  
2105 ATTGAGGCGCAAGAAAGTTGTCCTGAGACACTGGGCGGCTTCTTAATCTTACTCTTC 2164  
4375 ATTGAGGCGCAAG-AGTTGTCCTGAGACACTGGGCGGCTTCTTAATCTTACTCTTC 4433  
2165 TGTCCCGGACAACTGTCTCAGAGTCTGCTACTATTTCTGAGGGGCTCTTAAGAGGGGCA 2224  
4434 TGTCCCGGACAACTGTCTCAGAGTCTGCTACTA-TCGAGGGGCGTCC-FAAGAGCGGCA 4491  
2225 GTCACTAGATCTTTTCCAGCCCACTAAGTTATGAATGGGAGCTTATCTTTTCA 2284  
4492 GTCACTAGATCTTTTCCAGCCCACTAAGTTATG-ACGAGGAGCTTATCTTTTCA 4549  
2285 ATGCTTTTCTAATTAATGCTTAAGAGCCCACTAATCTTGTAGGAGAGCATTTAGCAA 2344  
4550 ATGCTTTTCTAATTAATGCTTAAGAGCCCACTAATCTTGTAGGAGAGCATTTAGCAA 4609  
2345 AAGCAGGGGCACTTAATCACTGAACATTAAGAGAGAAACCCGTTGTGTGTCCTG 2404  
4610 AAGCAGGGGCACTTAATCACTGAACATTAAGAGAGAAACCCGTTGTGTGTCCTG 4668  
2405 CTGAGAGAGAAATTAATCTGAGTCTGGGCAAGAGCAATTAATGAGCAGCAAA 2464  
4669 CTGAGAGAGAAATTAATCTGAGTCTGGGCAAGAGCAATTAATGAGCAG-CAAA 4727  
2465 GAATGCGCGCTCTGTTCAAGTTAACTAAGAGTTCACTTCTTCTTCAAAAGCA 2524  
4728 GAATGCGCGCTCTGTTCAAGTTAACTAAGAGTTCACTTCTTCTTCAAAAGCA 4787  
2525 GTACCCCTCAGACCCCAAGGCCCAAGAGACTCCAAAAGATTGTTAAGAGACTTAAGAGC 2584  
4788 GTACCCCTCAGACCCCAAGGCCCAAGAGACTCCAAAAGATTGTTAAGAGACTTAAGAGC 4847



Db	6976	AAACCTCAGGGATTCGATGATCTACTAGCTCGGGTAGAATCTTCAAGGGTTGGGCAAG	7035
Oy	4804	GCCTTCCCCTGATAGCAGAAAAAGGCCAAGAGTAAATAAAGGCACTAGTTCATGAATA	4863
Db	7036	GCCTTCCCCTGATAGCAGAAAAAGGCCAAGAGGATTAATAAGGCACTAGTTCATGAATA	7095
Oy	4864	ATTCGAGATTGGAATTCCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGGCC	4923
Db	7096	ATTCGAGATTGGAATTCCCCGAGGCTTACAGAGTGACAAATAGCCCTGCTTTCAGGCC	7155
Oy	4924	AAGATTAACCCAGGGATATCCCAAGCGTTAAGGTATATACATATACATTACATTGGGCTCGA	4983
Db	7156	AAGATTAACCCAGGGATATCCCAAGCGTTAAGGTATATACATATACATTACATTGGGCTCGA	7215
Oy	4984	AGGCAACAGTCTCTCAGGGAGAGGTGAGAAAAATGAATGAATAAATCAAGAGCACTCTAAA	5043
Db	7216	AGGCAACAGTCTCTCAGGGAGAGGTGAGAAAAATGAATGAATAAATCAAGAGCACTCTAAA	7275
Oy	5044	AAAGCAACCCAGGAAACCCACCTCACATAGCCCTGYTCTGTTCCTATAGCCTTAAAAAG	5103
Db	7276	AAAGCAACCCAGGAAACCCACCTCACATAGCCCTGYTCTGTTCCTATAGCCTTAAAAAG	7335
Oy	5104	ATCTGCACCTTCCCCCAAAAAGCAGAGACTTAGCCCATACGAAATGCTGTATAGGAAGGCC	5163
Db	7336	ATCTGCACCTTCCCCCAAAAAGCAGAGACTTAGCCCATACGAAATGCTGTATAGGAAGGCC	7395
Oy	5164	TTCATATACCAATGACCTTGTGCTTACCCAGACAGCCACCTTAGTTGACAGATCACT	5223
Db	7396	TTCATATACCAATGACCTTGTGCTTACCCAGACAGCCACCTTAGTTGACAGATCACT	7455
Oy	5224	CCTTAGCCAAATATACAGAGTTCTTAAAACTTACAAAGAACCTATCCTCGAGAGAG	5283
Db	7456	CCTTAGCCAAATATACAGAGTTCTTAAAACTTACAAAGAACCTATCCTCGAGAGAG	7515
Oy	5284	GAAAAGACTATTTCCACCCCMGTGACATGATATTGTCAATCCCTTCTCTTAATCCC	5343
Db	7516	GAAAAGACTATTTCCACCCCTGTGTGACATGATATTGTCAATCCCTTCTCTTAATCCC	7575
Oy	5344	CATCCCTGATATCATCTCTGGAGAGACCCCTACCCAGTATTTTAAATVTAACCCCACTGGG	5403
Db	7576	CATCCCTGATATCATCTCTGGAGAGACCCCTACCCAGTATTTTAAATVTAACCCCACTGGG	7635
Oy	5404	TTAAAGTGGCTGAGTGAAGTCTTGGAATACATCACACTTAGATCAATCCTGATATCTGC	5463
Db	7636	TTAAAGTGGCTGAGTGAAGTCTTGGAATACATCACACTTAGATCAATCCTGATATCTGC	7695
Oy	5464	CANAAGAACCTGAAAAATCAGAGACCAACGCTAGTATTCTGTGAACTCTTAAGGATT	5523
Db	7696	CANAAGAACCTGAAAAATCAGAGACCAACGCTAGTATTCTGTGAACTCTTAAGGATT	7755
Oy	5524	TGCGCTGCTCTTCAACAACAACAGAGAGGAAATTAATCTAAATCATTAATCCCCATG	5583
Db	7756	TGCGCTGCTCTTCAACAACAACAGAGAGGAAATTAATCTAAATCATTAATCCCCATG	7814
Oy	5584	GSCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTAAACCTSTTTTCACTCACTGCA	5643
Db	7815	GSCTCCCTTATCATATTTTCTCTKTAAGTGTSTTTAAACCTSTTTTCACTCACTGCA	7874
Oy	5644	CCCCCTCCATGCGCTGTATGACCAATGCTCCCTTACAAAGATTTCTATGAGAAATG	5703
Db	7875	CCCCCTCCATGCGCTGTATGACCAATGCTCCCTTACAAAGATTTCTATGAGAAATG	7934
Oy	5704	CAGGCTCCCGGAAATATTTGATGCCCATCGTATAGAGCTTTTSPAAAGGAACCCCAAC	5763
Db	7935	CAGGCTCCCGGAAATATTTGATGCCCATCGTATAGAGCTTTTSPAAAGGAACCCCAAC	7994
Oy	5764	TTCACTGCCCAACCCCATATAGCCCGCAACTGCTATCACTGTGCCACTCTTTGATGAT	5823
Db	7995	TTCACTGCCCAACCCCATATAGCCCGCAACTGCTATCACTGTGCCACTCTTTGATGAT	8054
Oy	5824	GCAATACTCATTTATTTGACAGGAAAAATGATTAATCTAGTTGTCTGAGAGACTTGA	5883

Db	8055	GCAAAATACCTATTAATGGAACGAAAAATGATTAATCTGATGTTGTCCTGGAGACCTTGGAA	8114
QY	5884	GTCACTGTCGTGTGGACTTATCTTACCCAACTGATATGTCGTGATGGGGGTGAGTTTCAA	5943
Db	8115	GTCACTGTCGTGTGGACTTATCTTACCCAACTGATATGTCGTGATGGGGGTGAGTTTCAA	8174
QY	5944	GATCAGGCGAAGAAAAACATGTAAAGAGTAATCTTCCCACTCACCCSGGGTTCATATGC	6003
Db	8175	GATCAGGCGAAGAAAAACATGTAAAGAGTAATCTTCCCACTCACCCGGGTTCATATGC	8234
QY	6004	ACCTCTAGCCCCCTTCAAAAGACATGATCTCTCAAAAATCAATAGAAACCCCTCCGATACCAT	6063
Db	8235	ACCTCTAGCCCCCTTCAAAAGACATGATCTCTCAAAAATCAATAGAAACCCCTCCGATACCAT	8294
QY	6064	ACTGCGCTGTGAACCTATTAAATACCAACCTCACTGGGGCTCATGAGATCTCGGCCCAA	6123
Db	8295	ACTGCGCTGTGAACCTATTAAATACCAACCTCACTGGGGCTCATGAGATCTCGGCCCAA	8354
QY	6124	AACCTTACTTAACTTGTGGATATGCTCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	6183
Db	8355	AACCTTACTTAACTTGTGGATATGCTCTCCCTGAACTTCARGCCATATGTTTCAATCCCT	8414
QY	6184	GTACTCTBAACAATGGAAACACTTCAGACAGAAATAAACAACAACCTTCGGTTTTAGTAGGA	6243
Db	8415	GTACTCTBAACAATGGAAACACTTCAGACAGAAATAAACAACAACCTTCGGTTTTAGTAGGA	8474
QY	6244	CCTCTGTGTTTCCATATGTGAAATTAACCCATACCTCAACCTCACTGTGTATAATTTAGC	6303
Db	8475	CCTCTGTGTTTCCATATGTGAAATTAACCCATACCTCAACCTCACTGTGTATAATTTAGC	8534
QY	6304	AATATCTACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCTCCCAACAATA	6363
Db	8535	AATATCTACATACACAAACCACTCCCAATGCATCAGGTGGGTAACTCTCCCAACAATA	8594
QY	6364	GTCGAGCCTACCCCTCAGGAATATTTTTGTGTGGTATCCTCAGGCTTATGCTTGTGTAAT	6423
Db	8595	GTCGAGCCTACCCCTCAGGAATATTTTTGTGTGGTATCCTCAGGCTTATGCTTGTGTAAT	8654
QY	6424	GGCCTTTCAAGATCTATATGTCTCTCTCAATTTCTAGTGGCCCCATATGRCATCTACACT	6483
Db	8655	GGCCTTTCAAGATCTATATGTCTCTCTCTCAATTTCTAGTGGCCCCATATGRCATCTACACT	8714
QY	6484	GAAACAATTTATACATATATGTCATCTTAAGCCCCGGCAACAAAAGATACCCATCTT	6543
Db	8715	GAAACAATTTATACATATATGTCATCTTAAGCCCCGGCAACAAAAGATACCCATCTT	8774
QY	6544	CCTTTGTATTAAGGAGAGGAGTGTCTAGGCACTAGGTCTGTGGCATTTGGCGGTATACA	6603
Db	8775	CCTTTGTATTAAGGAGAGGAGTGTCTAGGCACTAGGTCTGTGGCATTTGGCGGTATACA	8834
QY	6604	ACCTCTACTAGTTTCTACTCAAACTATCTCAAGAACTTAATGGGGGACATGGAACGGGTC	6663
Db	8835	ACCTCTACTAGTTTCTACTCAAACTATCTCAAGAACTTAATGGGGGACATGGAACGGGTC	8894
QY	6664	GCCGACTTCCTGTGACCTTGGCAAGATCAACTTAATCTCCTTACGACGATATGTCCTTGRA	6723
Db	8895	GCCGACTTCCTGTGACCTTGGCAAGATCAACTTAATCTCCTTACGACGATATGTCCTTCAA	8954
QY	6724	AATGGAAGACTTTTGAATGTCGTAAACCGCGAAGAGAGGGGGAAACCTGTTATTTTAAAGG	6783
Db	8955	AATGGAAGACTTTTGAATGTCGTAAACCGCGAAGAGAGGGGGAAACCTGTTATTTTAAAGG	9014
QY	6784	GAAAGATGCTGTTATTTATGTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATT	6843
Db	9015	GAAAGATGCTGTTATTTATGTTAAATCAATCCGGAATCGTCACTGAGAAAGTTTAAAGAAATT	9074
QY	6844	CSAGATGGAATACAGCTATAGAGAGAGAGCTTGGAAACACTGGAACCTTGGGGCTTCTC	6903
Db	9075	CSAGATGGAATACAGCTATAGAGAGAGAGCTTGGAAACACTGGAACCTTGGGGCTTCTC	9134
QY	6904	AGCCRAATGGAATGCCCGGAGATTCCTCCCTTTTAGAGCTCTTACAGAGTATTAATATGCTA	6963
Db	9135	AGCCRAATGGAATGCCCGGAGATTCCTCCCTTTTAGAGCTCTTACAGAGTATTAATATGCTA	9194



Db	Accession	Version	Source	Organism	Reference	Journal	Features
Oy	6964				TTCCCTCTTGGAGCCCTGTATCTTTTAACTCCCTGTGTAACCTTGTCTCTTCCAGATGGAA		
Db	9195				CTCCCTCTTGGAGCCCTGTATCTTTTAACTCCCTGTGTAACCTTGTCTCTTCCAGATGGAA		
Oy	7024				GCTGTAAATCTAACATATGGAGCCCAAGATGACATCCAAAGCACTAGATCTTACCGAGACCC		
Db	9255				GCTGTAAATCTAACATATGGAGCCCAAGATGACATCCAAAGCACTAGATCTTACCGAGACCC		
Oy	7084				CTGACCGGACCTGTATGACCCACGATCTGTATGTTAATGACATCAAAGGCACTCCCTCTGAG		
Db	9315				CTGACCGGACCTGTATGACCCACGATCTGTATGTTAATGACATCAAAGGCACTCCCTCTGAG		
Oy	7144				GAAATCTAGCTGACCAACCTCTACAGCCCAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG		
Db	9375				GAAATCTAGCTGACCAACCTCTCTACTACAGCCCAATTCAGAGAGAGAGAGAGAGAGAGAGAG		
Oy	7204				SGTCGGCCAACTCCCAAGCACTTATGATTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAG		
Db	9435				CGTGGCCAACTCCCAAGCACTTATGATTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAG		
Oy	7264				AGGATCTAGCTGATTTCTTCTAGGCTGATTTAAGATCTTTAAGCTTTAGGAGAGAGAGAGAG		
Db	9495				AGGATCTAGCTGATTTCTTCTAGGCTGATTTAAGATCTTTAAGCTTTAGGAGAGAGAGAGAG		
Oy	7324				ACATCCACCTTTAAACAGGGGGCTTGAACATAGTCAACCTGACCAATCAGAGAGAGAGAG		
Db	9555				ACATCCACCTTTAAACAGGGGGCTTGAACATAGTCAACCTGACCAATCAGAGAGAGAGAGAG		
Oy	7384				ACTAAATATGCTATTAAGGCAAG		
Db	9615				ACTAAATATGCTATTAAGGCAAG		
Oy	7444				AGAGCAGCAGCAG		
Db	9675				AGAGCAGCAGCAG		
Oy	7504				CAACCCCTTTGGGATCCCTCCCTTTGATGGAGAGCTGTGTTTCAATCTATTTTCACTCT		
Db	9735				CAACCCCTTTGGGATCCCTCCCTTTGATGGAGAGAGCTGTGTTTCAATCTATTTTCACTCT		
Oy	7564				ATTAAATCTTGCACCTGCR 7582		
Db	9795				ATTAAATCTTGCACCTGCA 9813		
RESULT 6	AY101585				10222 bp	DNA	linear PRI 11-FEB-2004
LOCUS	AY101585						
DEFINITION							
VERSION	AY101585.1						
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
PUBMED							
AUTHORS							
TITLE							
JOURNAL							
FEATURES							
source							

	/mol_type="Homo sapiens"
	/organism="Homo sapiens"
	/isolate="79"
	/isolation_source="PBMC of individual 79"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7q21-q22"
	/sex="male"
	/cell_type="PBMC"
	/note="caucasian
	isolated by PCR; ERVWE1 locus, allele B
	endogenous_virus= HERV-W"
LTR	1..780
mRNA	join(248..811,2887..9813)
mRNA	/note="putative mRNA transcript 1; gag-pol pseudogene"
mRNA	join(248..811,7585..9813)
mRNA	/product="envelope glycoprotein"
mRNA	/note="putative mRNA transcript 2"
CDS	join(248..811,9248..9813)
	/note="putative mRNA transcript 3; no coding region"
	7812..9428
	/note="syncytin"
	/codon_start=1
	/product="envelope glycoprotein"
	/protein_id="PAM68164.1"
	/db_xref="GI:37544406"
	/translation="MAlPFHIFLPYTLNLSFTLTAPPCRCMTSSSPDYQFLRMORP GNIDASYSRLSKGPTFPAHHMPRNCBSATLCHMAHTHYTGKINSCGGELAV TYWWTFTQMGSDGGVDQAERKEVKETISQLTFRVHGTSIPPEOMNFSTELNTTSV HMLVLFPMTTLGLHEVSADNPITNCMI CLELNRPVGSIPVEOMNFSTELNTTSV LVCLPLSNLETHTNSLTCVCNKSTNYTTNSOCIRWTPTOICVCPSGI FPGGTSA YRLNSSSESBCFLSLFVPMTITYEODLSYISKRNRPVLI LFPVGAGALG TGIGGITSTGFYKLSDSELNDMERVAUSLVTLQDUNSLAAVTLDNRALDLTR E RGCTCFLTEBCCCYVNOSGIVTEKXKLRDR IQRABEELRNGLPWGLSOMPWLILP FLPLAAILLLLFGPCIFNLVNFVSSRIEA VKLOMEPRMQSKTIYRRPLDRPAS P RSVDNIKGTPPEIRISAQPI LRPNSGSR"
	9487..10222
ORIGIN	
Query Match	85.0%; Score 6448; DB 9; Length 10222;
Best Local Similarity	95.7%; Pred. No. 0;
Matches 6681; Conservative 199; Mismatches 50; Indels 53; Gaps 27;	
OY	606 GATGGGAAACGTTCCCGCAAGACAAAAAGCCCCCTTAAGACGATTTCGGARAAATTGGGA 665
Db	2886 GATGGGAAAAGTTCCCGCAAGACAAAAAGCCCCCTTAAGACGATTTCGGAGAATTGGGA 2945
OY	666 MCAATTGACCCTCAGACATAAGAAAGAACGACTATATTCTTCGACAGTGCCTG 725
Db	2946 CCAATTGACCCTCAGACATAAGAAAGAACGACTATATTCTTCGACAGTGCCTG 3005
OY	726 GCACTCCGAGGGAAGTAAATTAACAACAATCTTAGAGTAGACYCTTTGTGAA 785
Db	3006 GCACTCCGAGGGAAGTAAATTAACAACAATCTTAGAGTAGACCTTTGTGAAA 3065
OY	786 AAGCAAATGAGTAGAAGTGCATAGTACAAATCTTTCACTTAAGACAACCTCAC 845
Db	3066 AAGCAAATGAGTAGAAGTGCATAGTACAAATCTTTCACTTAAGACAACCTCAC 3128
OY	846 AATTATGTAAGAGTGAATTATAGCCCTACAGGAACCTTCAGAGTCTACCTCCATAC 905
Db	3126 AATTATGTAAGAGTGAATTATAGCCCTACAGGAACCTTCAGAGTCTACCTCCATAC 3185
OY	906 CCAAGCATCCCAGACTCTCTTCCCAATTAAGACACCCCTTCAACCCAATGCTCAA 965
Db	3186 CCAAGCATCCCAGACTCTTCCCAATTAAGACACCCCTTCAACCCAATGCTCAA 3245
OY	966 AAGAGATTAACAAGAGGTAAACAGTGAACCAAGAGTCCCAATATCCCAATTATGA 1025
Db	3246 AAGAGATTAACAAGAGGTAAACAGTGAACCAAGAGTCCCAATATATCCCAATTATGA 3305
OY	1026 CCCCTCCAGAGTGGGAAGAAATTCGGGCCAGAGAGTGCATGTCCTTTTYY 1085

Db 3306 CCCCT-CCAGCATGGGAGAGAAATTGCGCCCAAGCCAGATGTCCTTTTC 3364  
Qy 1086 TCTCCCAAGCTTTAAAGCAAAATAAAAAGAGCTTAGGTAATTTCTGATTAATCTGATAG 1145  
Db 3365 TCTCCCAAGCTTTAAAGCAAAATAAAAAGAGCTTAGGTAATTTCTGATTAATCTGATAG 3424  
Qy 1146 CTATTTTBTGTTTACAAGGGTTAGACAAATTTCTTGTATCTGACATGAGAGATATATA 1205  
Db 3425 CTATTTTBTGTTTACAAGGGTTAGACAAATTTCTTGTATCTGACATGAGAGATATATA-A 3483  
Qy 1206 TGTCACTGCTAAATCAGACACTAACCCCAATGAGAGAGTGCACCACTTAATGAGGCT 1265  
Db 3484 TGTCACTGCTAAATCAGACACTAACCCCAATGAGAGAGTGCACCACTTAATGAGGCT 3543  
Qy 1266 GAGGTTTGGCGATCTGTGTATCTCAGTCAGTCATGAGATNGAGATGACAAACAGAAAG 1325  
Db 3544 GAGGTTTGGCGATCTGTGTATCTCAGTCAGTCATGAGATNGAGATGACAAACAGAAAG 3600  
Qy 1326 AAAAGAAATGATTTCCCAAGGCCAGACAGCATTTCCAGTCASACCTCATTTGGGAC 1385  
Db 3601 AAAAGAAATGATTTCCCAAGGCCAGACAGCATTTCCAGTCATGACCTCATTTGGG--- 3657  
Qy 1386 ACAGAAATCAGTAATCAGTGGAGATTTGGCTGAGACATTTGCTTAATCTTGTGCTAATA 1445  
Db 3658 ACAGAAATCAGTAATCAGTGGAGATTTGGCTGAGACATTTGCTTAATCTTGTGCTAATA 3716  
Qy 1446 GGAATTAAGAAATCAGTGGAGATTTGGCTGAGACATTTGCTTAATCTTGTGCTAATA 1505  
Db 3717 GGAATTAAGAAATCAGTGGAGATTTGGCTGAGACATTTGCTTAATCTTGTGCTAATA 3775  
Qy 1506 GGGAGAGGAAAGAAATCTCTGCTGCTTTCTGGAGAGACTTAAGGAGCATTTGAGAAAGC 1565  
Db 3776 -GGGAGAGGAAAGAAATCTCTGCTGCTTTCTGGAGAGACTTAAGGAGCATTTGAGAAAGC 3834  
Qy 1566 GTGCTCTCTGTCACTGACTCTTCTGAAAGGCAACTATCTTAAGGCTAAGTTTATCA 1625  
Db 3835 GTGCTCTCTGTCACTGACTCTTCTGAAAGGCAACTATCTTAAGGCTAAGTTTATCA 3894  
Qy 1626 CTCAGTCACTGACAGACATTAAG-AAAAAATCTCAAAAGTCTGCCGTAGGCCGAGCAAA 1684  
Db 3895 CTCAGTCACTGACAGACATTAAG-AAAAAATCTCAAAAGTCTGCCGTAGGCCGAGCAAA 3954  
Qy 1685 ACTTAAGAAACCTTAATGAACTTGGCACTGTGGTTTTTATTAATGAGATCAGAGAGAC 1744  
Db 3955 ACTTAAGAAACCTTAATGAACTTGGCACTGTGGTTTTTATTAATGAGATCAGAGAGAC 4014  
Qy 1745 AGGCGGAAACAGGACAAACGGGATTTAAAAAAGGCAACGCTTATGATGATGACCTCAGG 1804  
Db 4015 AGGCGGAAACAGGACAAACGGGATTTAAAAAAGGCAACGCTTATGATGATGACCTCAGG 4074  
Qy 1805 CAAAGTGAATTTGAGGCTCTGAGAAAGGAAAGCTGGGCAAAATGGAATGCTAATAG 1864  
Db 4075 CAAAGTGAATTTGAGGCTCTGAGAAAGGAAAGCTGGGCAAAATGGAATGCTAATAG 4134  
Qy 1865 GCTTGCTTCCAGTGGCTTACAAGACACTTTAAAAAAGTGTCCAAAGTGAAGTAAAG 1924  
Db 4135 GCTTGCTTCCAGTGGCTTACAAGACACTTTAAAAAAGTGTCCAAAGTGAAGTAAAG 4194  
Qy 1925 CCGGCCCCCTTGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCAAG 1984  
Db 4195 CCGGCCCCCTTGTCCATGCCCCCTTATTTCAAGGGAATCACTGGAAGGCCCACTGCCCAAG 4254  
Qy 1985 GGAAGAAAGTCTTTGAGTGAAGACACTTAACAGATGATCCAGAGCAGAGAGCTGAGAG 2044  
Db 4255 GGAAGAAAGTCTTTGAGTGAAGACACTTAACAGATGATCCAGAGCAGAGAGCTGAGAG 4314  
Qy 2045 TGCCTGGGAGAGCGCATCCCATGCACTACCTTCAAGAGCCCTGGGATATGCTTAAAC 2104  
Db 4315 TGCCTGGGAGAGCGCATCCCATGCACTACCTTCAAGAGCCCTGGGATATGCTTAAAC 4374  
Qy 2105 ATTGAAGGCGAGAGAGTGTCTCTGAGACATGCTGGGCTTCTTAAATGCTTCTTC 2164  
Db 4375 ATTGAAGGCGAGAG-AGGTTGTCTCTGAGACATGCTGGGCTTCTTAAATGCTTCTTC 4433  
Qy 2165 TGTCCCGGACACTGTCTCTCCAGATCTGTCACTATTTGAGGGGGTCCNTAAGACGGACA 2224  
Db 4434 TGTCCCGGACAACTGTCTCTCCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGACA 4491  
Qy 2225 GTCACTAGATCTTTTTCACGCACTTAAGTTATGAACTGGGAGCTTTATTTCTTTCAC 2284  
Db 4492 GTCACTAGATCTTTTTCACGCACTTAAGTTATGAACTGGGAGCTTTATTTCTTTCAC 4549  
Qy 2285 ATGCTTTCTTAATGATGTTGAAGCCCACTACTGTTTGAAGGAGACATTTCAAGAA 2344  
Db 4550 ATGCTTTCTTAATGATGTTGAAGCCCACTACTGTTTGAAGGAGACATTTCAAGAA 4609  
Qy 2345 AAGCAGGGGCGCATTAACACTGAACTAGAGAGAGAACACCGGTTTGTGTCCTG 2404  
Db 4610 AAGCAGGGGCGCATTAACACTGAACTAGAGAGAGAACACCGGTTTGTGTCCTG 4668  
Qy 2405 CTTGAGAGAGAAATTAATCTGAAAGTCTGGGCAACAGAGACAAATATGACGAGCCAAA 2464  
Db 4669 CTTGAGAGAGAAATTAATCTGAAAGTCTGGGCAACAGAGACAAATATGACGAG-CAAA 4727  
Qy 2465 GAATGCCGTCCTGTCAAGTTAACTTAAGGATTTCACTTCTTCCCTACCAAGGCA 2524  
Db 4728 GAATGCCGTCCTGTCAAGTTAACTTAAGGATTTCACTTCTTCCCTACCAAGGCA 4787  
Qy 2525 GTACCCCTCAGACCCCAAGGCCCAAGAGATTCAAAAGATGTTGAAGACTTAAAGC 2584  
Db 4788 GTACCCCTCAGACCCCAAGGCCCAAGAGATTCAAAAGATGTTGAAGACTTAAAGC 4847  
Qy 2585 CCAAGGCTTGAATTAACATGATGATTAATCTCCCTGCAATTAATCCGTAGTGAAGAGG 2644  
Db 4848 CCAAGGCTTGAATTAACATGATGATTAATCCCTGCAATTAATCCGTAGTGAAGAGG 4900  
Qy 2645 CACAGAAACCAAGGACAGTGGAGGGTGTGTCAGATCTCAGATTAATCAATGAGAGC 2704  
Db 4901 TACAGAAACCAAGGACAGTGGAGGGTGTGTCAGATCTCAGATTAATCAATGAGAGC 4958  
Qy 2705 GGTGTCTCTTAATCCCAAGCTGATCTAGCCCTTAATCTGTGATGATTTCCAAATACAGA 2764  
Db 4959 GGTGTCTCTTAATCCCAAGCTGATCTAGCCCTTAATCTGTGATGATTTCCAAATACAGA 5018  
Qy 2765 GGAAGCAGATGCTTACABTCTCTGACCTTGAAGATGCTTCTTCTGCAATCCTGATCA 2824  
Db 5019 GGAAGCAGATGCTTACABTCTCTGACCTTGAAGATGCTTCTTCTGCAATCCTGATCA 5078  
Qy 2825 TCTGACTCTCAATCTGTTGCTTGAAGATTAATCTTCAAAACCAATCTCAACTCAC 2884  
Db 5079 TCTGACTCTCAATCTGTTGCTTGAAGATTAATCTTCAAAACCAATCTCAACTCAC 5138  
Qy 2885 CTGACTTCTTAAACCCCAAGGTTCAAGGATGAGTCCCATCTATTTGGCCAGGACTTAAAGC 2944  
Db 5139 CTGACTTCTTAAACCCCAAGGTTCAAGGATGAGTCCCATCTATTTGGCCAGGACTTAAAGC 5198  
Qy 2945 CCAAGACTTGAATCAATTAATCAATCTGACACTCTTGTCTTCTGATGATGATTT 3004  
Db 5199 CCAAGACTTGAATCAATTAATCAATCTGACACTCTTGTCTTCTGATGATGATTT 5256  
Qy 3005 ACTTTTGGCGCCATTCAGAAACCTTGGGCACTGAAGCAACCAACRCCTTMAATTT 3064  
Db 5257 ACTTTTGGCGCCATTCAGAAACCTTGGGCACTGAAGCAACCAACRCCTTMAATTT 5316  
Qy 3065 CCTGCGTACTGTGGCTACAGGTTTCAAAAGGATGAGTCCATCTGCTCAGAGAGT 3124  
Db 5317 CCTGCGTACTGTGGCTACAGGTTTCAAAAGGATGAGTCCATCTGCTCAGAGAGT 5376  
Qy 3125 TAAATCTTGAATTAATTAATCAAGGCAACAGGCTCTGATGAGATTAATCA 3184  
Db 5377 ----TACTTAAGGCTTAAATTAATCAAGGCAACAGGCTCTGATGAGATTAATCA 5432  
Qy 3185 GCTTAATAGGCTTATCTCATCTCAAAACCTTAAGGCACTTAAGGATTTCTTGGCT 3244  
Db 5433 GCTTAATAGGCTTATCTCATCTCAAAACCTTAAGGCACTTAAGGATTTCTTGGCT 5492



```
Db 7636 TTAAGTGGCTGAGTGGAGCTTGGATACATCACACTTGAATCTGAATCTGGATCTGC 7695
Qy 5464 CAAAGAAACCTGAAAATCCAGAGACAAAGCTAGCTATTCCTGTGAACCTCTAGAGATT 5523
Db 7696 CAAAGAAACCTGAAAATCCAGAGACAAAGCTAGCTATTCCTGTGAACCTCTAGAGATT 7755
Qy 5524 TGGCCCTGCTCTTCAACACACACAGAGAGAAAGTAACTAAATCATTAATCCCCATG 5583
Db 7756 TGGCCCTGCTCTTCAACACACACAGAGAGAAAGTAACTAAATCATTAAT-CCCCATG 7814
Qy 5584 GSCCTCCCTTATCATATTTTCTCTKASTGTSTTTTAAACCTSTTTTACTCTCACTGCA 5643
Db 7815 GCCCTCCCTTATCATATTTTCTCTTACTGTCTTTTAAACCTCTTCACTCTCACTGCA 7874
Qy 5644 CCCCCCTCATSCCCTGTATGACAGTAGCTCCCTTACMAAGATTCTATGAGAAATG 5703
Db 7875 CCCCCCTCATSCCCTGTATGACAGTAGCTCCCTTACMAAGATTCTATGAGAAATG 7934
Qy 5704 CAGCGTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAAACCCCAAC 5763
Db 7935 CAGCGTCCCGGAAATATTGATGCCCATCGTATAGAGTCTTTTAAAGGAAACCCCAAC 7994
Qy 5764 TTCACTGCGCACACCCATATGCGCGCACTGCTATCACTTGCCACTTTTGCATGCAAT 5823
Db 7995 TTCACTGCGCACACCCATATGCGCGCACTGCTATCACTTGCCACTTTTGCATGCAAT 8054
Qy 5824 GCAAAATCTCATTTTGGACAGAGAAATGATTAATCCTAGTTGCTTGAAGAGACTTGA 5883
Db 8055 GCAAAATCTCATTTTGGACAGAGAAATGATTAATCCTAGTTGCTTGAAGAGACTTGA 8114
Qy 5884 GTCACTGCTGTGGACTTACTTCAACCAACCTGTATGCTGAATGGGGGTGAGTTCAA 5943
Db 8115 GTCACTGCTGTGGACTTACTTCAACCAACCTGTATGCTGAATGGGGGTGAGTTCAA 8174
Qy 5944 GATCAGCGAAGAGAAAAATGTTAAAGAAATTTCTCCAACTCAACCGGGGTACATGAC 6003
Db 8175 GATCAGCGAAGAGAAAAATGTTAAAGAAATTTCTCCAACTCAACCGGGGTACATGAC 8234
Qy 6004 ACCCTGAGCCGCTCAAAAGACTAGATGCTCAAACTACATGAAACCTCCGATACCAT 6063
Db 8235 ACCCTGAGCCGCTCAAAAGACTAGATGCTCAAACTACATGAAACCTCCGATACCAT 8294
Qy 6064 ACTGCGCTGTAGACTTATTAATCAACCTCACTGCGCTCACTGAGGTTGCGGCCAA 6123
Db 8295 ACTGCGCTGTAGACTTATTAATCAACCTCACTGCGCTCACTGAGGTTGCGGCCAA 8354
Qy 6124 AACCTACTAATGTTGATATGCTCCCTGAACTTCAAGCCATATGTTCAATCCCT 6183
Db 8355 AACCTACTAATGTTGATATGCTCCCTGAACTTCAAGCCATATGTTCAATCCCT 8414
Qy 6184 GTAACCTAAACAAATGGAACAACTTACAGACAGAAATTAACACACCTCCGTTTATGTA 6243
Db 8415 GTAACCTAAACAAATGGAACAACTTACAGACAGAAATTAACACACCTCCGTTTATGTA 8474
Qy 6244 CCTCTTGTTCCTAATSTGAAATTAACCATTAACCTCAAACTCACTGTGTAAATTTAGC 6303
Db 8475 CCTCTTGTTCCTAATSTGAAATTAACCATTAACCTCAAACTCACTGTGTAAATTTAGC 8534
Qy 6304 AATACTACATACACAAACCTCCCAATGATCAAGTGGGTAACTCCTCCACACAAATA 6363
Db 8535 AATACTACATACACAAACCTCCCAATGATCAAGTGGGTAACTCCTCCACACAAATA 8594
Qy 6364 GTCTGCTACCTCCTAGAGAAATTTTGTCTGTGTGATCTCAAGCTTATCGTTTGAAT 6423
Db 8595 GTCTGCTACCTCCTAGAGAAATTTTGTCTGTGTGATCTCAAGCTTATCGTTTGAAT 8654
Qy 6424 GGCCTCTCAGAAATCTATGTGCTCTCTCATTTCTAGTGCCCGATAGCATCTACACT 6483
Db 8655 GGCCTCTCAGAAATCTATGTGCTCTCTCATTTCTAGTGCCCGATAGCATCTACACT 8714
Qy 6484 GAAACAAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAGATGCCATCTT 6543
Db 8715 GAAACAAGATTATACAGTTATGTCATATCTAAGCCCGCAACAAAGATGCCATCTT 8774
Qy 6544 CCTTTGTATAGAGCAGAGTGTCTAGTGACATAGGTACTGCACTTGGCGGTATCACA 6603
Db 8775 CCTTTGTATAGAGCAGAGTGTCTAGTGACATAGGTACTGCACTTGGCGGTATCACA 8834
Qy 6604 ACTCTACTCAGATTCTATCTACAAATCTATCTCAAGAACTTAAATGGGACATGGAACGGGTC 6663
Db 8835 ACTCTACTCAGATTCTATCTACAAATCTATCTCAAGAACTTAAATGGGACATGGAACGGGTC 8894
Qy 6664 GCGGACCTCGGTGACCTTGAAGATCACTTAACCTCCCTAGACAGATGTCCTTCRA 6723
Db 8895 GCGGACCTCGGTGACCTTGAAGATCACTTAACCTCCCTAGACAGATGTCCTTCRA 8954
Qy 6724 AATCGAAGACTTTTAACTGCTTAAACCGCTGABAGAGGGGAACTGTATTTATTTAAGG 6783
Db 8955 AATCGAAGACTTTTAACTGCTTAAACCGCTGABAGAGGGGAACTGTATTTTATTTAAGG 9014
Qy 6784 GAAAGATGCTGTATTTATTTATTAATCAATCCGGAATGCTCACTGAGAAAGTTAAGAAAT 6843
Db 9015 GAAAGATGCTGTATTTATTTATTTAATCAATCCGGAATGCTCACTGAGAAAGTTAAGAAAT 9074
Qy 6844 CSAGATGGAATTAACAACGTAAAGAGAGAGACTTCCAAACACTGACACCTGGGCGCTCCTC 6903
Db 9075 CSAGATGGAATTAACAACGTAAAGAGAGAGACTTCCAAACACTGACACCTGGGCGCTCCTC 9134
Qy 6904 AGCCBATGATGCGCTGAGATTTCTCCCTCTTGAAGACCTCTAGACGCTATTAATTTGCTA 6963
Db 9135 AGCCBATGATGCGCTGAGATTTCTCCCTCTTGAAGACCTCTAGACGCTATTAATTTGCTA 9194
Qy 6964 CTCCTCTTTGAGACCTGTATCTTTTAACTTCTCTTGTAACTTTGTCTTTCAGAAATGAA 7023
Db 9195 CTCCTCTTTGAGACCTGTATCTTTTAACTTCTCTTGTAACTTTGTCTTTCAGAAATGAA 9254
Qy 7024 GCTGTAAATTAACAATGAGAGCCCAAGATGACATCCCAAGACTTAAGTACCCGAGACC 7083
Db 9255 GCTGTAAATTAACAATGAGAGCCCAAGATGACATCCCAAGACTTAAGTACCCGAGACC 9314
Qy 7084 CTGAGACCGGCTGTAGCCCAAGATGATGTTAATGATCACTCAAGGACACCCCTCTGAG 7143
Db 9315 CTGAGACCGGCTGTAGCCCAAGATGATGTTAATGATCACTCAAGGACACCCCTCTGAG 9374
Qy 7144 GAAATCTGAGCTGACACCTTACTACGCCCCAATTCAGCAGAGACGATTAGAGGCT 7203
Db 9375 GAAATCTGAGCTGACACCTTACTACGCCCCAATTCAGCAGAGACGATTAGAGGCT 9434
Qy 7204 SGTGCGCAACCTCCCAACAGACATTAAGTTTCCGTTGAGATGGGGGCTGAGAGAC 7263
Db 9435 SGTGCGCAACCTCCCAACAGACATTAAGTTTCCGTTGAGATGGGGGCTGAGAGAC 9494
Qy 7264 AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCCTAAGCTAGSTGGGAAAGTGACC 7323
Db 9495 AGGACTAGCTGATTTCTTAGGCTGATTAAGAAATCCCTAAGCTAGSTGGGAAAGTGACC 9554
Qy 7324 ACATCAACCTTTAAACAGGGGCTTGCACTTAGCTCAACCTGACCAATCAGAGAGCTC 7383
Db 9555 ACATCAACCTTTAAACAGGGGCTTGCACTTAGCTCAACCTGACCAATCAGAGAGCTC 9614
Qy 7384 ACTAAATGCTAATTTAGGCAAGACAGAGAGTAAAGAAATAGCCATCAATTTATGCGTG 7443
Db 9615 ACTAAATGCTAATTTAGGCAAGACAGAGAGTAAAGAAATAGCCATCAATTTATGCGTG 9674
Qy 7444 AGAGCAAGCAGAGAGGACATGATGCGGATATTAACCCCAAGTGTGAGCCCGCAACGG 7503
Db 9675 AGAGCAAGCAGAGAGGACATGATGCGGATATTAACCCCAAGTGTGAGCCCGCAACGG 9734
Qy 7504 CAAACCCCTTTGGGCTCCCTCCCTTTGTATGGAGGCTGTGTTCATGCTATTTCACTCT 7563
Db 9735 CAAACCCCTTTGGGCTCCCTCCCTTTGTATGGAGGCTGTGTTCATGCTATTTCACTCT 9794
Qy 7564 ATTAAATCTTGACATGCTGR 7582
Db 9795 ATTAAATCTTGACATGCTGA 9813
```

RESULT 7  
AX329572 . 56093 bp DNA linear PAT 09-JAN-2002  
LOCUS Sequence 81 from Patent WO0194629.  
ACCESSION AX329572  
VERSION AX329572.1 GI:18102550  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
AUTHORS 1  
TITLE Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endres, G.,  
Horrigan, S., Soppet, D.R. and Weaver, Z.  
Cancer gene determination and-therapeutic screening using signature  
gene sets  
JOURNAL Patent: WO 0194629-A 81 13-DEC-2001;  
FEATURES Avalon Pharmaceuticals (US)  
SOURCE 1..56093  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 85.0%; Score 6448; DB 6; Length 56093;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;  
QY 606 GATGGGAAACGTTCCCGCAGACAAAACGCCCTTAGAGCTATTCGGAAAAATTGGGA 665  
DB 30953 GATGGGAAACGTTCCCGCAGACAAAACGCCCTTAGAGCTATTCGGAAAAATTGGGA 31012  
QY 666 MCAATTTGACCCCTGAGACCTAAGAAAAGACCTTATCTCTGAGTGGCCGCTG 725  
DB 31013 CCAATTTACCCCTGAGACCTAAGAAAAGACCTTATCTCTGAGTGGCCGCTG 31072  
QY 726 GCACTCTGAGGGAAGTAAATTATTAACACCATCTTACAGCTAGACCTTTTGTAGA 785  
DB 31073 GCATCTCCGAGGGAAGTAAATTATTAACACCATCTTACAGCTAGACCTTTTGTAGA 31132  
QY 786 AAGGCAATGGAAGTAAAGTCCATTAAGCAAACTTTCTTTCAATTAAGACAACTCAC 845  
DB 31133 AAGGCAATGGAAGTAAAGTCCATTAAGCAAACTTTCTTTCAATTAAGACAACTCAC 31192  
QY 846 AATTATGTAAGAAAGTGAATTAATGCTCCAGAGAAAGCTTCAAGTCTAAGCTCCCTATC 905  
DB 31193 AATTATGTAAGAAAGTGAATTAATGCTCCAGAGAAAGCTTCAAGTCTAAGCTCCCTATC 31252  
QY 906 CCAGCATCCCGCAGCTCTTCCCCAMTATTAAGAACCCCTTCAACCCAAATGCTCCAA 965  
DB 31253 CCAGCATCCCGCAGCTCTTCCCCAMTATTAAGAACCCCTTCAACCCAAATGCTCCAA 31312  
QY 966 AAGGAGATGACAAAAGGTTAAACAGTGAAACCAAAGAGTCCCAATTTCCCAATTATGA 1025  
DB 31313 AAGGAGATGACAAAAGGTTAAACAGTGAAACCAAAGAGTCCCAATTTCCCAATTATGA 31372  
QY 1026 CCCCTCCCAAGAGGAGGAGAAAGAAATGGGCCAGCAGAGTCAAGTGCCTTTTTC 1085  
DB 31373 CCCCT-CCAAAGAGGAGGAGAAAGAAATGGGCCAGCAGAGTCAAGTGCCTTTTTC 31431  
QY 1086 TCTCCAGACTTAAAGCAAAATTAAGACCTTAGTAATTTCTAGATTAATCTGATG 1145  
DB 31432 TCTCCAGACTTAAAGCAAAATTAAGACCTTAGTAATTTCTAGATTAATCTGATG 31491  
QY 1146 CTATATTTGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAA 1205  
DB 31492 CTATATTTGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAA 31550  
QY 1206 TGTGCTGCTAAATAGACACTTAACCCCAATGAGAGAGTGCACCATTAATGAGAGCT 1265  
DB 31551 TGTGCTGCTAAATAGACACTTAACCCCAATGAGAGAGTGCACCATTAATGAGAGCT 31610

QY 1266 GAGGTTTGGGATCTCTGTTATCTCACTGAGCTCAATGATTAAGATGACAAAGAG 1325  
DB 31611 GAGGTTTGGGATCTCTGTTATCTCACTGAGCTCAATGATTAAGATGACAAAGAG 31667  
QY 1326 AAAGANATGATTTCCCAAGAGCCAGCAGGAGTCCCAAGTCTASACCTCATTTGGGAG 1385  
DB 31668 AAAGANATGATTTCCCAAGAGCCAGCAGGAGTCCCAAGTCTASACCTCATTTGGG 31724  
QY 1386 ACAGAAATCAATGAGAGATTTGCTGAGACATTTGCTAATCTTGTGTCTAATA 1445  
DB 31725 ACAGAGATCAAGACAT-AGAGATTGGTGTGAGACATTTGCTAATCTTGTGTCTAATA 31783  
QY 1446 GGACTAAGGAAATCTAAGAGAAATCTAATTAATTAATTAATTAATTAATTAATTA 1505  
DB 31784 GGACTAAGGAAATCTAAGAGAAATCTAATTAATTAATTAATTAATTAATTAATTA 31842  
QY 1506 GGGGAGGAAAGAAATCTCTGCTCTTCTGAGAGACTAAGGAGGATTTGAGAGAG 1565  
DB 31843 -GGGAGGAGAGAAATCTCTGCTCTTCTGAGAGACTAAGGAGGATTTGAGAGAG 31901  
QY 1566 GTGCTCTCTGCTCACTGACTCTTCTGAGAGCCAACTAATCTTAAAGGTAATTA 1625  
DB 31902 GTGCTCTCTGCTCACTGACTCTTCTGAGAGCCAACTAATCTTAAAGGTAATTA 31961  
QY 1626 CTGAGTCAAGTGCAGACATTAG-AAAGAACTTCAAAAGCTCCGCTAGCCCGAGCA 1684  
DB 31962 CTGAGTCAAGTGCAGACATTAGAAAGAAAGCTCCGCTAGCCCGAGCA 32021  
QY 1685 ACTTAGAAACCTTATGAACTTTGCAACCTTGTGCTTTTAAATTAAGATCAGAGAGAG 1744  
DB 32022 ACTTAGAAACCTTATGAACTTTGCAACCTTGTGCTTTTAAATTAAGATCAGAGAGAG 32081  
QY 1745 AGGCGAAACAGAGCAAAAGGATTAAGAAAGGCAAGCTTTATGATGATGACCT 1804  
DB 32082 AGGCGAAACAGAGCAAAAGGATTAAGAAAGGCAAGCTTTATGATGATGACCT 32141  
QY 1805 CAAGTGAATTTGAGAGCTCTGAGAAAGGAAAGGAAATTTGAATGCTTAATAG 1864  
DB 32142 CAAGTGAATTTGAGAGCTCTGAGAAAGGAAAGGAAATTTGAATGCTTAATAG 32201  
QY 1865 GCTTGCTTCCAGAGGAGCTCAAGAGGACCTTTAAAGAAAGTTCAGAGTGAAGTAA 1924  
DB 32202 GCTTGCTTCCAGAGGAGCTCAAGAGGACCTTTAAAGAAAGTTCAGAGTGAAGTAA 32261  
QY 1925 CCGCCCTCTGCTCAATGCTCCCTTATTTCAAGAGAAATCACTGAAAGGCCACTG 1984  
DB 32262 CCGCCCTCTGCTCAATGCTCCCTTATTTCAAGAGAAATCACTGAAAGGCCACTG 32321  
QY 1985 GGAACAAAGTCTTTTGAAGTCAAGAGCCATTAACAGATGATTCAGACAGAGCTGAG 2044  
DB 32322 GGAACAAAGTCTCTGAGTCAAGAGCCATTAACAGATGATTCAGACAGAGCTGAG 32381  
QY 2045 TGCCCTGGGGGAGGAGGCAATCCCATGACATCAAGCCCTCAAGAGCCCTGGGATG 2104  
DB 32382 TGCCCTGGGGGAGGAGGCAATCCCATGACATCAAGCCCTCAAGAGCCCTGGGATG 32441  
QY 2105 ATTGAGGAGCAGAAAGTGTCTCTGAGACATGAGTGGGCTTTCTTATGCTTACT 2164  
DB 32442 ATTGAGGAGCAGAG-AGGTTGTCTCTGAGACATGAGTGGGCTTTCTTATGCTTACT 32500  
QY 2165 TGTCCCGGACAACTGTCTCTCAAGTCTGTCACTATTCTGAGAGGGTCCNTAAGAGG 2224  
DB 32501 TGTCCCGGACAACTGTCTCTCAAGTCTGTCACTATTCTGAGAGGGTCCNTAAGAGG 32558  
QY 2225 GTCACTAGATCTTTTCCAGACCTAAGTATTAAGATCTGGGAGCTTATCTTTTAC 2284  
DB 32559 GTCACTAGATC-TTCTCCAGACCTAAGTATTAAGATCTGGGAGCTTATCTTTTAC 32616  
QY 2285 ATGCTTTTCTAATTAATGTTGAAAGCCCACTAATCTTGTGAGAGAGACATTTAG 2344  
DB 32617 ATGCTTTTCTAATTAATGTTGAAAGCCCACTAATCTTGTGAGAGAGACATTTAG 32676



```
QY 2345 AAGAGGGGCAATTATACCTTGAACATAGAGAAAGAAACACCGTTTGTGTTCCCTG 2404
DB 32677 AAGCAGGGGGCAATTATACACTGGAACATAGAGAAAGAAACACCGTTTGTGTTG - CCCCG 32725
QY 2405 CTTGAGGAAGGAATTAATTCCTGAAGCTGGGCAACGAAGACAAATATGAGAGCCAAA 2464
DB 32726 CTTGAGGAAGGAATTAATTCCTGAAGCTGGGCAACGAAGACAAATATGAGAGAG - CAAA 32724
QY 2465 GAATGCCGCTCTGTTCAAGTTAACTAAAGATTCCACTTCTTCCCTACCAAGCA 2524
DB 32795 GAATGCCGCTCTGTTCAAGTTAACTAAAGATTCCACTTCTTCCCTACCAAGCA 32854
QY 2525 GTACCCCTCAGACCCCAAGGCCCAACAGGATTCCAAAAGATTGTTAAGACTTAAAGC 2584
DB 32855 GTACCCCTCAGACCCCAAGGCCCAACAGGACTCCAAAAGATTGTTAAGACTTAAAGC 32914
QY 2585 CCAGGCTTAGTAAACCATGATTAACCTCCCTGCAATATTCGGTAGTGGATTGAGAGG 2644
DB 32915 CCAGGCTTAGTAAACCATGATTAACCTCCCTGCAATATTC-----AATTTAAGAG 32967
QY 2645 CACAGAAACCCAGTGAAGTGAAGGTTAGTCAAGATCTCAGGATTATCAATGAGAGC 2704
DB 32968 TACAGAAACCCCAAGCAAGTGAAGGTTAGTCAAGATCTCAGGATTATCAAT - GAGGC 33025
QY 2705 GGTGTCTTTTATACCAAGCTGTACTAGCCCTTATCTGTGTTTCCCAATACAGA 2764
DB 33026 TGTGTCTCTATATAGCAGGTGTACTAGCCCTTATCTGTGTTTCCCAATACAGA 33085
QY 2765 GGAAGCAGAGTGTATTAACSTCTGAGACTTMAAGATGCTTCTGTCGCTGTACA 2824
DB 33086 GGAAGCAGAGTGTATTAACSTCTGAGACTTMAAGATGCTTCTGTCGCTGTACA 33145
QY 2825 TCCTGACTCTCAATTCTGTTTGGCTTGAAGATCTTCAACCCAGCATCTCACTAC 2884
DB 33146 TCCTGACTCTCAATTCTGTTTGGCTTGAAGATCTTCAACCCAGCATCTCACTAC 33205
QY 2885 CTGACATTTTTTACCCCAAGGGTTCAGAGATAGYCCCATCTATTTTGGCCAGGCA 2944
DB 33206 CTGACATTTTTTACCCCAAGGGTTCAGAGATAGYCCCATCTATTTTGGCCAGGCA 33265
QY 2945 CCAAGACTTGAAGYCAATTTTCACTTCTGACACTTCTTCTGCTGCTGCTGCTG 3004
DB 33266 CCAAGACTTGAAGYCAATTTTCACTTCTGACACTTCTTCTGCTGCTGCTGCTG 33323
QY 3005 ACTTTGCGYGCCTTTCAGAAACCTTGGCATCAAGCCCAAGGCTCTTMAATT 3064
DB 33324 ACTTTGCGYGCCTTTCAGAAACCTTGGCATCAAGCCCAAGGCTCTTMAATT 33383
QY 3065 CCTGCGYACCTGTGCTACAMGGTTTCCAAACATGCTGCTGCTGCTGCTGCTG 3124
DB 33384 CCTGCGYACCTGTGCTACAMGGTTTCCAAACATGCTGCTGCTGCTGCTGCTG 33443
QY 3125 TAAATCTAGGCTAATATTTTCCAAAGGCAACGAGGCTCTGCTGCTGCTGCTG 3184
DB 33444 TAAATCTAGGCTAATATTTTCCAAAGGCAACGAGGCTCTGCTGCTGCTGCTG 33499
QY 3185 GCGTATCTGCTTATCTCTCATCTCAATTCCTAAAGCACTAAGGRTTCTTGGGRT 3244
DB 33500 GCGTATCTGCTTATCTCTCATCTCAATTCCTAAAGCACTAAGGRTTCTTGGGRT 33559
QY 3245 AAVAGYTTCTGCGAAMATGATTTCCAGGTTTGGCRAATATGCGAGYCATTAATA 3304
DB 33560 AAVAGYTTCTGCGAAMATGATTTCCAGGTTTGGCRAATATGCGAGYCATTAATA 33618
QY 3305 CAAATTAATTAAGAACTCAGAAAGCAATACCATTTATATAGATGATYAMCTGAGYR 3364
DB 33619 CAAATTAATTAAGAACTCAGAAAGCAATACCATTTATATAGATGATYAMCTGAGYR 33678
QY 3365 AAGTGGCTTTCAGAGCCCTTAAAGAGGCTTAAACCAAGYCCAGGTTAAAGTTGCC 3424
DB 33679 AAGTGGCTTTCAGAGCC-----TAAACCAAGCCCAAGGTTAAAGTTGCC 33724
QY 3425 AACGGGCAAGACTTTTSTYATATRTCAAGAAAAAACAGRAAYAGCTCTGGAGTCC 3484
DB 33725 AACGGGCAAGACTTTTSTYATATRTCAAG - AAAAAACAGRAATAGCTCTAGAGTCC 33783
QY 3485 TTACACAGTCCRAGGAYGAGCTTGCAACCTRTGCGYACTGASTAAAGAAATGATG 3544
DB 33784 TTACACAGTCCRAGGAYGAGCTTGCAACCTRTGCGYACTGASTAAAGAAATGATG 33843
QY 3545 TAGTGGCAAGAGGTTGACCTTATGTTTAAAGGATGAGTGGCACTAGAGYKATAGT 3604
DB 33844 TAGTGGCAAGAGGTTGACCTTATGTTTAAAGGATGAGTGGCACTAGAGYKATAGT 33903
QY 3605 CTGAAGCATTAAATTAATACAGGAGAGATCTTACTGTGACATCTCATGAGTGA 3664
DB 33904 CTGAAGCATTAAATTAATACAGGAGAGATCTTACTGTGACATCTCATGAGTGA 33963
QY 3665 AYRGCTACTCACTGTCTAAAGAGACTTGTGGCTGTGACACACTGTTTACTTAAATRT 3724
DB 33964 AYRGCTACTCACTGTCTAAAGAGACTTGTGGCTGTGACACACTGTTTACTTAAATRT 34023
QY 3725 AGGCTCTATTCTGAARGGCCAGTGTGACCTGTGCACTTGTGCAACTCTTAACCCAG 3784
DB 34024 AGGCTCTATTCTGAARGGCCAGTGTGACCTGTGCACTTGTGCAACTCTTAACCCAG 34083
QY 3785 YCNCAATTTCTTCAAGACATGAAAGAAAGATARAAYTACTGTCAACATTAATTTCT 3844
DB 34084 YCNCAATTTCTTCAAGACATGAAAGAAAGATARAAYTACTGTCAACATTAATTTCT 34143
QY 3845 AAACCTATGCTGAGAGGGGACTTGTAGAGTTCCYTTGACTGATCTGACCTTCAAC 3904
DB 34144 AAACCTATGCTGAGAGGGGACTTGTAGAGTTCCYTTGACTGATCTGACCTTCAAC 34202
QY 3905 TTGTATCTGATGGAAGTCTTGTGAAAAGAAAGAAAGCTGAAAAGGGGTTATGCACT 3964
DB 34203 TTGTATCTGATGGAAGTCTTGTGAAAAGAAAGAAAGCTGAAAAGGGGTTATGCACT 34262
QY 3965 GTCAGATTAATGGAATATTTGAAAGTATCCCTCACTCCAGAACTAGTCTGAGCTR 4024
DB 34263 GTCAGATTAATGGAATATTTGAAAGTATCCCTCACTCCAGAACTAGTCTGAGCTR 34322
QY 4025 GCAGAACTAATAGCTTCATYTKGGGCACTAGAAATTAGAGAAAGGAAAGGAAATTA 4084
DB 34323 GCAGAACTAATAGCTTCATYTKGGGCACTAGAAATTAGAGAAAGGAAAGGAAATTA 34382
QY 4085 TATACAGCTCTATATGCTGACCTGATGCTGATGCTGATGCTGATGCTGATGCTG 4144
DB 34383 TATACAGCTCTATATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 34442
QY 4145 AGGGAATTCCTAATCTTGAAGGAACTATCAATCAATCAAGAAAGCACTTGAATTA 4204
DB 34443 AGGGAATTCCTAATCTTGAAGGAACTATCAATCAATCAAGAAAGCACTTGAATTA 34502
QY 4205 TTATTTGCTGATCAAGAACTTAAAGAGTGTGATCTTCACTGCTGCTGCTGCTGCTG 4264
DB 34503 TTATTTGCTGATCAAGAACTTAAAGAGTGTGATCTTCACTGCTGCTGCTGCTGCTG 34562
QY 4265 GGAAGGAAAGGGGAAATTAAGAAATGGAAGTGAAGCAATATTTGAAGCAAAAGCTGCA 4324
DB 34563 GGAAGGAAAGGGGAAATTAAGAAATGGAAGTGAAGCAATATTTGAAGCAAAAGCTGCA 34622
QY 4325 AGGCAAGACCTCTCATTAAGAAATGCTTATTAATCTTCACTGATAGAGTATCCCTTC 4384
DB 34623 AGGCAAGACCTCTCATTAAGAAATGCTTATTAATCTTCACTGATAGAGTATCCCTTC 34682
QY 4385 CGGGAACCAAGCCCACTGATCTGAGGAGGAAACAGATAGGGGAACCTCAAGAG - CAG 4443
DB 34683 CGGGAACCAAGCCCACTGATCTGAGGAGGAAACAGATAGGGGAACCTCAAGAG - CAG 34742
QY 4444 TTTTCTGCTGCGGAGCGTTAGCACTGAAGAGGAAATATTTTGTGCTGCAACTAT 4503
DB 34743 TTTTCTGCTGCGGAGCGTTAGCACTGAAGAGGAAATATTTTGTGCTGCAACTAT 34802
QY 4504 CCAATGGAATTAATTAACCTTCACTCACTTCACTTCACTTCACTTCACTTCACTTCA 4563
```

Db 34803 CCAATGGAATTAATCTTAAACCTTCAATCAACCTTCACTTAGCATCATGACACCA 34862  
Qy 4564 TCATATGGCCAAATCATATTTATTTACTGACACAGCCCTTTCAAACCTATCAAGCABATAT 4623  
Db 34863 TCAGATGGCCAAATCATATTTATTTACTGACACAGCCCTTTCAAACCTATCAAGCABATAT 34922  
Qy 4624 CAGGCGCTGTGAATGTGGCCBARAATAATCCCTGCTATATGCGCAAGCTCCTTAGG 4683  
Db 34923 CAGGCGCTGTGAATGTGGCCBARAATAATCCCTGCTATATGCGCAAGCTCCTTAGG 34982  
Qy 4684 ABAACAABAACAGGCCATTAACCTGTBARAARACTGGCACTGATTTTACCAACAGGCC 4743  
Db 34983 AGAACAABAACAGGCCATTAACCTGTBARAARACTGGCACTGATTTTACCAACAGGCC 35042  
Qy 4744 AAACCTCAGGATTTCAAGTATCTATGCTGGTAAATATTTTCAAGGTTGGGCABAG 4803  
Db 35043 AAACCTCAGGATTTCAAGTATCTATGCTGGTAAATATTTTCAAGGTTGGGCABAG 35102  
Qy 4804 GCCTTCCCTGTAGAGACAGAAAGGCCAAGGTAATAAGGCACTAGTTCATGAATA 4863  
Db 35103 GCCTTCCCTGTAGAGACAGAAAGGCCAAGGTAATAAGGCACTAGTTCATGAATA 35162  
Qy 4864 ATTCCAGATTCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTCCAGGCC 4923  
Db 35163 ATTCCAGATTCGACTTCCCGAGGCTTACAGAGTGAACAATAGCCCTGCTTCCAGGCC 35222  
Qy 4924 ACAGTAACCCAGGAGATCCAGAGCGTTAGATATACATATCACTTACATGCGCCTGA 4983  
Db 35223 ACAGTAACCCAGGAGATCCAGAGCGTTAGATATACATATCACTTACATGCGCCTGA 35282  
Qy 4984 AGGCCACAGTCTCTAGGGAAGGTGAGAAATGAAATGAATAAATCAAGGCACTCTAAA 5043  
Db 35283 AGGCCACAGTCTCTAGGGAAGGTGAGAAATGAAATGAATAAATCAAGGCACTCTAAA 35342  
Qy 5044 AAGCAAAACCCAGAAACCCACTCAATGCGCTGTGCTAGCTATAGCCTTAAAAAGA 5103  
Db 35343 AAGCAAAACCCAGAAACCCACTCAATGCGCTGTGCTAGCTATAGCCTTAAAAAGA 35402  
Qy 5104 ATCTGCAACTTTCCCAAAAGACAGAGACTTAGGCCATACGAATGCTGTATGAAAGGCC 5163  
Db 35403 ATCTGCAACTTTCCCAAAAGACAGAGACTTAGGCCATACGAATGCTGTATGAAAGGCC 35462  
Qy 5164 TTCTAATCAATGACCTTGTGCTTGAACCAAGACAGCCCACTTAGTTCAGACATCACT 5223  
Db 35463 TTCTAATCAATGACCTTGTGCTTGAACCAAGACAGCCCACTTAGTTCAGACATCACT 35522  
Qy 5224 CCTTAGCCAAATATCAACAAGTCTTAAACATTAACAAGAACCTATCCCTGAGAAAGG 5283  
Db 35523 CCTTAGCCAAATATCAACAAGTCTTAAACATTAACAAGAACCTATCCCTGAGAAAGG 35582  
Qy 5284 GAAAGAAGACTATTCACCCMWGTGACATGTATATGTCAGTCCCTTCTCTAATTTCC 5343  
Db 35583 GAAAGAAGACTATTCACCCMWGTGACATGTATATGTCAGTCCCTTCTCTAATTTCC 35642  
Qy 5344 CATCCCTAGATACATCTGCGAAGAACCTTACCCGATTTTAATTTAATCCCAATTGGG 5403  
Db 35643 CATCCCTAGATACATCTGCGAAGAACCTTACCCGATTTTAATTTAATCCCAATTGGG 35702  
Qy 5404 TTAAAGTGGCTGAGTGGAGTCTTGATATCATCACTTGAATCAAACTCTGATACTGC 5463  
Db 35703 TTAAAGTGGCTGAGTGGAGTCTTGATATCATCACTTGAATCAAACTCTGATACTGC 35762  
Qy 5464 CAAAGGAACCTGAAAATCCAGAGACAAACGCTAGTATTTCTGTGAACCTCTAGAGGAT 5523  
Db 35763 CAAAGGAACCTGAAAATCCAGAGACAAACGCTAGTATTTCTGTGAACCTCTAGAGGAT 35822  
Qy 5524 TGGCGCTGCTCTTCAAACAACAACAGAGGAAATTAATCAATAAATCCCCCATG 5583  
Db 35823 TGGCGCTGCTCTTCAAACAACAACAGAGGAAATTAATCAATAAATCCCCCATG 35881  
Qy 5584 GSCCTCCCTATCATATTTTCTCTKTASTGTSTTTTACCTSTTTCACTCTCACTGA 5643  
Db 35882 GSCCTCCCTATCATATTTTCTCTKTACTGTTCTTTTACCTCTTTCACTCTCACTGA 35941

Qy 5644 CCCCTCAATGCGCGTATGACCAAGTATGCCCTTACCMAGATTTCTATGAGATG 5703  
Db 35942 CCCCTCAATGCGCGTATGACCAAGTATGCCCTTACCMAGATTTCTATGAGATG 36001  
Qy 5704 CAGCGTCCCGAAATTTGATGCGCATGTATAGAGTCTTTSTAAGGAACCCCAAC 5763  
Db 36002 CAGCGTCCCGAAATTTGATGCGCATGTATAGAGTCTTTSTAAGGAACCCCAAC 36061  
Qy 5764 TTCACTGCGCAACCCATATGCCCCCAACTGCTATACATGCGCACTCTTGGATGAT 5823  
Db 36062 TTCACTGCGCAACCCATATGCCCCCAACTGCTATACATGCGCACTCTTGGATGAT 36121  
Qy 5824 GCAATATCTATTAATGACAGGAAATATTAATCTATGTTGCTCTGAGAGACTTGA 5883  
Db 36122 GCAATATCTATTAATGACAGGAAATATTAATCTATGTTGCTCTGAGAGACTTGA 36181  
Qy 5884 GTCACTGTCTGTGACTTATCTTACCCAACTGTATGTCTGATGCGGGTGAATTCAA 5943  
Db 36182 GTCACTGTCTGTGACTTATCTTACCCAACTGTATGTCTGATGCGGGTGAATTCAA 36241  
Qy 5944 GATCAGGCAAGAAAAACATTTAAAGAAATATCTCCCACTCACCGGGTATGAGC 6003  
Db 36242 GATCAGGCAAGAAAAACATTTAAAGAAATATCTCCCACTCACCGGGTATGAGC 36301  
Qy 6004 ACCCTAGCCCTTACAAGGACTAGATCTCTCAAACTACATGAAACCTTCCGTATCCAT 6063  
Db 36302 ACCCTAGCCCTTACAAGGACTAGATCTCTCAAACTACATGAAACCTTCCGTATCCAT 36361  
Qy 6064 ACTGCGCTGTGAAGCTTATTAATACCACTTCACTGCGCTCCATGAGGTCTCGGCCAA 6123  
Db 36362 ACTGCGCTGTGAAGCTTATTAATACCACTTCACTGCGCTCCATGAGGTCTCGGCCAA 36421  
Qy 6124 AACCTTACTAATCTGTGGAATATGCTCCCTGAGCTTCAAGGCAATGTTCAATCCCT 6183  
Db 36422 AACCTTACTAATCTGTGGAATATGCTCCCTGAGCTTCAAGGCAATGTTCAATCCCT 36481  
Qy 6184 GTACCTGAACAGGAACAACCTTCAAGCAAGAAATTAACAACCACTTCCGTTTATGAGA 6243  
Db 36482 GTACCTGAACAGGAACAACCTTCAAGCAAGAAATTAACAACCACTTCCGTTTATGAGA 36541  
Qy 6244 CCTCTGTCTTCAATSTGAAATTAACCATATCTCAAACTCACTGCTGTATTAATGAC 6303  
Db 36542 CCTCTGTCTTCAATSTGAAATTAACCATATCTCAAACTCACTGCTGTATTAATGAC 36601  
Qy 6304 AATATCTATACACACCAACTCCCATGCAATGATGAGTGGGTATCTCTCCCAACAATA 6363  
Db 36602 AATATCTATACACCAACTCCCATGCAATGATGAGTGGGTATCTCTCCCAACAATA 36661  
Qy 6364 GTCTGCTTACCTCAGAAATATTTTGTCTGTGCTGATCTGACCTATCGTGTGTAAT 6423  
Db 36662 GTCTGCTTACCTCAGAAATATTTTGTCTGTGCTGATCTGACCTATCGTGTGTAAT 36721  
Qy 6424 GGCCTTCAAGATCTATGTGCTTCTCTCATTTCTAGTCCCTCATGRCATCTTACCT 6483  
Db 36722 GGCCTTCAAGATCTATGTGCTTCTCTCATTTCTAGTCCCTCATGRCATCTTACCT 36781  
Qy 6484 GAACAGATTTATCAAGTATGTATATCTTAAGCCCGGCAACAAGAGTCCCACTTCT 6543  
Db 36782 GAACAGATTTATCAAGTATGTATATCTTAAGCCCGGCAACAAGAGTCCCACTTCT 36841  
Qy 6544 CCTTTGTATAGAGAGAGAGTCTAGTATCTAGTATCTGCACTTGGCGGTATCA 6603  
Db 36842 CCTTTGTATAGAGAGAGAGTCTAGTATCTAGTATCTGCACTTGGCGGTATCA 36901  
Qy 6604 ACCTTACTCAGTCTTATCAAACTATCTCAAGACTAATTTGGGACATGGAACGGGTC 6663  
Db 36902 ACCTTACTCAGTCTTATCAAACTATCTCAAGACTAATTTGGGACATGGAACGGGTC 36961  
Qy 6664 GCGGACTCCTGTGCTACTTGGACATCACTTAACTTCCCTTACAGCAGTATGCTTGA 6723  
Db 36962 GCGGACTCCTGTGCTACTTGGACATCACTTAACTTCCCTTACAGCAGTATGCTTGA 37021

OY		6724	AATCGAAGACCTTTAACTGCTGAACCCGTGTAPAAAGGGGGAACCTGTTTATTGTTTGGG	6783
Db		37022	AATCGAAGACTTTTAACCTTGCTAACCCGTGAAGAAGGGGGAACTGTTTATTGTTTGGG	3708
OY		6784	GAAAGATGCTGTTATTATGTATTAATCAATCCGGAATCGTCAGTAGMAAGTTTAAAGAAAT	6843
Db		37082	GAAAGATGCTGTTATTATGTATTAATCAATCCGGAATCGTCAGTAGMAAGTTTAAAGAAAT	3714
OY		6844	CSAGATCGAATACAACGTAACAGCAGARAGCTTCGAAACTGGAACCTCGGGGCTCTCTC	6903
Db		37142	CSAGATCGAATACAACGTAACAGCAGARAGCTTCGAAACCTGGAACCTCGGGGCTCTCTC	3720
OY		6904	AGCCBAGATGATCCCGTGAATTCGCCCTCTTAGAGACTCTAGAGAGCATATATTTGCTA	6963
Db		37202	AGCCBAGATGATCCCGTGAATTCGCCCTCTTAGAGACTCTAGAGAGCATATATTTGCTA	3726
OY		6964	CTCCTCTTTTGAACCCGTGTATCTTTTACCCTCTGTTTAACTTGTCTCTCCAGATGAA	7023
Db		37262	CTCCTCTTTTGAACCCGTGTATCTTTTACCCTCTGTTTAACTTGTCTCTCCAGATGAA	3732
OY		7024	GCTGTAAACTTCAAATGAGGCCCAAGATGCAGTCCAAAGATAAGATCTACCGAGACC	7083
Db		37322	GCTGTAAACTTCAAATGAGGCCCAAGATGCAGTCCAAAGATAAGATCTACCGAGACC	3738
OY		7084	CTGAGCGGGCCGYTAGGCCAAGATCTGATGTTTATGACATCAAGAACCCCTCTCTGAG	7143
Db		37382	CTGAGCGGGCCGYTAGGCCAAGATCTGATGTTTATGACATCAAGAACCCCTCTCTGAG	3744
OY		7144	GAAATCTGAGCTGACCAACTCTTACTAGGCCCAATCAGCAGGAACAGTTAGAGCGGT	7203
Db		37442	GAAATCTGAGCTGACCAACTCTTACTAGGCCCAATCAGCAGGAACAGTTAGAGCGGT	3750
OY		7204	SGTCGGCCAACCTCCCAACAGCACTTAGTTTTCTGTTGAGATGGGGGACTGAGAGAC	7263
Db		37502	SGTCGGCCAACCTCCCAACAGCACTTAGTTTTCTGTTGAGATGGGGGACTGAGAGAC	3756
OY		7264	AGGATACCTGAGTTTCTTAGCTGATGTAAAGATCCCTTAAGCTAGTGGGAAGTGACCC	7323
Db		37562	AGGATACCTGAGTTTCTTAGCTGATGTAAAGATCCCTTAAGCTAGTGGGAAGTGACCC	3762
OY		7324	ACATCAACCTTTTAAACAGGGGGCTTGCAACTTAGTVCACACTGACCATCAGAGAGCTC	7383
Db		37622	ACATCAACCTTTTAAACAGGGGGCTTGCAACTTAGTVCACACTGACCATCAGAGAGCTC	3768
OY		7384	ACTAAATGCTAATTAGGCAAGAAGAGGTAAAGAAATAGCAATCATYTATTCMTG	7443
Db		37682	ACTAAATGCTAATTAGGCAAGAAGAGGTAAAGAAATAGCAATCATYTATTCMTG	3744
OY		7444	AGAGCACAGCAGAGGAGCAATGATCGGATATTAACCAAGTYTTCAGCCGCAACGG	7503
Db		37742	AGAGCACAGCAGAGGAGCAATGATCGGATATTAACCAAGTYTTCAGCCGCAACGG	3780
OY		7504	CAACCCCCCTTTGGGGCCCTCCCTTGTATGGGAGCTGTGTTTCAATGCTATTTCACTCT	7563
Db		37802	CAACCCCCCTTTGGGGCCCTCCCTTGTATGGGAGCTGTGTTTCAATGCTATTTCACTCT	3786
OY		7564	ATTAAATCTTGCAACTGCR	7582
Db		37862	ATTAAATCTTGCAACTGCA	37880
RESULT 8				
HSAC000064				
LOCUS	HSAC000064	56093 bp	DNA	linear
ACCESSION	Human BAC clone RG08M05 from 7q21-7q22, complete sequence.			
VERSION	AC000064.1	GI:1669369		
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1 (bases 1 to 56093)			

AUTHORS  
TITLE  
JOURNAL  
REFERENCES  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Paulley/A.  
The sequence of H. sapiens BAC clone RG083M05  
Unpublished (1996)  
2 (bases 1 to 56093)  
Mammals D

Department of Genetics, Washington University  
St. Louis, MO 63108, USA  
e-mail: sapiens@watson.wustl.edu

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

SOURCE INFORMATION: This clone is from the first release of the human BAC library. The library contains cloned DNA from a human male fibroblast cell line 9788K. For references see: Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et al., Genomics 34:213-218 (1996).

VECTOR: pBEO

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:  
The orientation of this clone is unknown. Actual start of this clone is at base position 1 of H\_RG083M05; actual end is at 560933 of H\_RG083M05

## Source

This clone contains STS SWSS1725.

1. .56093

```
/organism="Homo sapiens"  
/mol type="genomic DNA"
```

```
/db_xref="taxon:9606"  
/chromosome="7"
```

```

/ map="7q21-7q22"
/ cytoband="7p21.31-7p21.32"

```

```
/clone_lib="CITB-978SK-E
```

```
complement(838. .1131)
/rpt family="ALU"
```

<1360. .16971  
/gene="WUGSC:H RG083M05

```
join(<1360. .1503,4181.
0493 0547 11531 1177
```

5485. .5347, 11631. .117  
14885. .14988, 16349. .16

```
/gene="MUGSC:H_RG083M05.  
/note="ATPase; strong si
```

biosynthesis protein PAS  
human CDNA C04279 (NID:6

```

/codon_start=1
/exact=1 id=AAAACACAC 111

```

```
/protein_id="AA046346.1"
/db_xref="GI:1669371"
```

```
/translation="KRLENIQKTL
SPDAQRCEILCNVIKNKLDGIN
```

SRQSI STREKVLTTLD FQKALRG  
TIO LPAKVCLKKEKYP E L E A N L P T

KPELLSKYIGASEQAVRDI FIRRA  
ITOI DWEETI ACUMIT AATSPB

QMLHSFLVSRLEILNVLSDSLPLA

MLLSKMS ELLPDESKFNMYRLXFG  
GKDQLFSQPVLRTASQEGCQELTT

KTRLAISQSHLMTALGHTRPSISE  
complement (4948. .5130)

```
/rpt_family="ALU"
```

**repeat\_region**

```
complement(4948. .5130)
/rpc family="ALU"
```

```
repeat_region complement(6581, .7133)
/rpt_family="L1"
repeat_region complement(7767, .8037)
/rpt_family="ALU"
repeat_region complement(8186, .8472)
/rpt_family="ALU"
misc_feature 8473..8625
/gene="WUSC.H.RG083M05.1"
/notes="match to human 3' EST H75782 (NID:g1049794), bases
287-444"
misc_feature 8841..9161
/gene="WUSC.H.RG083M05.1"
/notes="match to human 5' EST H75921 (NID:g1050050), bases
21-348"
misc_feature 9481..9547
/gene="WUSC.H.RG083M05.1"
/notes="match to human 5' EST N2627 (NID:g1130501), bases
276-343"
repeat_region complement(12612, .12907)
/rpt_family="ALU"
misc_feature 13670..13793
/gene="WUSC.H.RG083M05.1"
/notes="match to human 5' EST H41382 (NID:917434), bases
143-266"
repeat_region 13794..13877
/rpt_family="ALU"
misc_feature 13878..13906
/gene="WUSC.H.RG083M05.1"
/notes="match to human 5' EST H41382 (NID:g917434), bases
30-58"
repeat_region 13907..14104
/rpt_family="ALU"
repeat_region complement(14110, .14137)
/rpt_family="L1"
repeat_region complement(15618, .15907)
/rpt_family="ALU"
repeat_region 17227..17522
/rpt_family="ALU"
misc_feature 18667..19235
/notes="match to human fetal brain 5' EST D61494
(NID:g970409), bases 1-255, and to human 3' EST R07476
(NID:g759399)"
repeat_region 19550..19670
/rpt_family="ALU"
misc_feature 21507..37303
/notes="similarity to various SS-RNA virus polypeptides;
pseudogene, region of matches and close matches to
multiple human ESTs, see R68740 (NID:g842257)"
exon 37316..37489
/notes="Grali prediction, score = 80"
/evidence="not experimental"
repeat_region complement(38938, .39224)
/rpt_family="ALU"
misc_feature 39225..39707
/notes="match to multiple human ESTs, see N30113
(NID:g1148633)"
repeat_region 39800..40085
/rpt_family="ALU"
repeat_region complement(40347, .40538)
/rpt_family="ALU"
repeat_region complement(40632, .40924)
/rpt_family="ALU"
repeat_region complement(42283, .42891)
/rpt_family="ALU"
repeat_region complement(45474, .45613)
/rpt_family="ALU"
misc_feature complement(45614, .45737)
/notes="match to human 3' EST H48898 (NID:g988738), bases
129-333"
misc_feature complement(46107, .47026)
/notes="match to multiple human ESTs, see N81064
(NID:g1243765), H48897 (NID:g988737), and M78831
(NID:g273146)"
```

```
repeat_region complement(47027, .47318)
/rpt_family="ALU"
misc_feature complement(47365, .47782)
/notes="match to multiple human ESTs, see M77495
(NID:g1319089)"
misc_feature 47898..48115
/notes="match to human 5' EST H62306 (NID:g1015138), bases
93-368"
repeat_region complement(48116, .48405)
/rpt_family="ALU"
misc_feature complement(48406, .48584)
/notes="match to human 3' EST N29952 (NID:g1148472), bases
250-455, and 5' EST R12730 (NID:g765806)"
repeat_region complement(48787, .49405)
/rpt_family="ALU"
misc_feature complement(49406, .49534)
/notes="match to human 3' EST R65794 (NID:g838432), bases
309-440"
repeat_region complement(49638, .49672)
/rpt_family="ALU"
misc_feature complement(49674, .49890)
/notes="match to human 3' EST N29952 (NID:g1148472) and 5'
EST N29938 (NID:g1148458), sequences are from opposite
ends of the same clone"
gene complement(49698, .51806)
/gene="WUSC.H.RG083M05.2"
complement(49698, .49888, 51575, .51806)
/notes="WUSC.H.RG083M05.2"
/notes="coded for by human cDNA M37389 (NID:g1319205),
R65891 (NID:g838529), R65794 (NID:g838432) and R65794
(NID:g838432)"
CDS /protein_id="BAB46345.1"
/CD_xref="GI:1669370"
/translat="MFYFQCGIIFCFQGVYQIGDVSVYDEQDKRYRQIRGF
IDQYCEKSAALTMLIPILSSPDQDPASIIIGPEDLPFRKVEYLFVCHAPSEYRK
SRSPPTVPTREKGYIWHVGPPTALITIKESVANHL"
complement(51576, .51758)
/exon complement(51576, .51758)
/gene="WUSC.H.RG083M05.2"
/notes="Grali prediction, score = 86"
/evidence="not experimental"
repeat_region complement(52052, .52329)
/rpt_family="L1"
misc_feature 55557..55843
Query Match 85.0%; Score 6448; DB 9; Length 56093;
Best Local Similarity 95.7%; Pred. NO. 0;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;
```

OY	966	AAAGAGATTAACA	AAAGGTTAAACA	GATGAACCAAGATGGCCAAATATTTCCCAATATTAGA	1025	
Db	31313	AAAGAGATTAACA	AAAGGTTAAACA	GATGAACCAAGATGGCCAAATATTTCCCAATATTAGA	31372	
OY	1026	CCCTCCCAAGCA	AGTGGAGGAAGAATTTGGCCACCAAGAGCATGTCTTTTXY	1085		
Db	31373	CCCTC	-CCAA	GCACTGGGAGGAAGAATTTGGCCACCAAGAGCATGTCTTTTTC	31431	
OY	1086	TCCTCCAGACTTAA	AGCAATTA	AAACAGACTTAAAGTAAATTTCTCAGATTAATCTGATAG	1145	
Db	31432	TCCTCCAGACTTAA	AGCAATTA	AAACAGACTTAAAGTAAATTTCTCAGATTAATCTGATAG	31491	
OY	1146	CTATATTTGTTT	TTTACAAGGTT	TAGACAAATTTCTTATCTGACATGAGAGATATATA	1205	
Db	31492	CTATATTTGTTT	TTTACAAGGTT	TAGACAAATTTCTTATCTGACATGAGAGATATATA	31550	
OY	1206	TGTCACCTGTCAA	ATCAGACATA	CCCCCAATAGAGAGAGTGGCCCACTTAATCTGACGCT	1265	
Db	31551	TGTCACCTGTCAA	ATCAGACATA	CCCCCAATAGAGAGAGTGGCCCACTTAATCTGACGCT	31610	
OY	1266	GAGRGTTTGGCGAT	CTCTGGTATCTCAGTCAGTCATATGATANGATGACA	CAAGAG	1325	
Db	31611	GAGRGTTTGGCGAT	CTCTGGTATCTCAGTCAGTCATATGATANGATGACA	CAAGAG	31667	
OY	1326	AAAGAAATGATTT	CCCCACAGGCCACAGACGATTTCCCACTTATGGGAC	1385		
Db	31668	AAAGAAATGATTT	CCCCACAGGCCACAGACGATTTCCCACTTATGGGAC	1385		
OY	1386	ACAGAAATCATGTA	CATGGGAGATTTGGTGCAGACATTTGGTAACTTGTGTGCTASA	1445		
Db	31725	ACAGAAATCATGTA	CATGGGAGATTTGGTGCAGACATTTGGTAACTTGTGTGCTASA	1445		
OY	1446	GGACTTAAGGAAA	CTASGAAGAAATCTAYGAATTA	CTCAATGATGCCATTAACA	1505	
Db	31784	GGACTTAAGGAAA	CTASGAAGAAATCTAYGAATTA	CTCAATGATGCCATTAACA	31842	
OY	1506	GGGGAAGGAGAAA	ATCTTAATCTGCTTTCTGAGAGACATAAGGAGCATTTGAGAAAGC	1565		
Db	31843	-GGGAAGGAGAAA	ATCTTAATCTGCTTTCTGAGAGAGCATTAAGGAGCATTTGAGAAAGC	31901		
OY	1566	GTGCGCTCTGTA	CACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGGTTAATTATCA	1625		
Db	31902	GTGCGCTCTGTA	CACCTGACTCTTCTGAAGGCCAACTAATCTTAAAGGTTAATTATCA	31961		
OY	1626	CTCAGTCAGCTG	CAGACATTAG-AAAAA	CTTCAAAAGTCGCGGTAGCCCGAGCAAA	1684	
Db	31962	CTCAGTCAGCTG	CAGACATTAG-AAAAA	CTTCAAAAGTCGCGGTAGCCCGAGCAAA	32021	
OY	1685	ACTTAGAAA	CCCTATTGA	CTTGGCAATCTCGTTTTTATATATAGATCAGAGAGAC	1744	
Db	32022	ACTTAGAAA	CCCTATTGA	CTTGGCAATCTCGTTTTTATATATAGATCAGAGAGAC	32081	
OY	1745	AGCGGGAACA	CGAACAAGGGATTA	AAAAAGCCCGTTATAGTACACCCCTCAG	1804	
Db	32082	AGCGGGAACA	CGAACAAGGGATTA	AAAAAGCCCGTTATAGTACACCCCTCAG	32141	
OY	1805	CAAGTGA	CTTTGAGGCTCTGAAAAAGGAAAAAGCTGGCAAA	TTGATGCTTAATAG	1864	
Db	32142	CAAGTGA	CTTTGAGGCTCTGAAAAAGGAAAAAGCTGGCAAA	TTGATGCTTAATAG	32201	
OY	1865	GCTTGCTTCA	GTGCGGTCTTACAAGACA	CTTTA	AAAAAGTTGCTCAAGTGAAG	1924
Db	32202	GCTTGCTTCA	GTGCGGTCTTACAAGACA	CTTTA	AAAAAGTTGCTCAAGTGAAG	32261
OY	1925	CGGCCCCCTT	CTGTCAATGCCCTTA	TTTCAAGGGAATCATGTGAAGCCCATGCCCCCAG	1984	
Db	32262	CGGCCCCCTT	CTGTCAATGCCCTTA	TTTCAAGGGAATCATGTGAAGCCCATGCCCCCAG	32321	
OY	1985	GGACAAAGGCT	TTTTGAGTCAAGAACCACTAA	CAGATGATCCAGACAGACAGACTGAAGG	2044	
Db	32322	GGACAAAGGCT	TTTTGAGTCAAGAACCACTAA	CAGATGATCCAGACAGACAGACTGAAGG	32381	
OY	2045	TGCCTTGGGCG	ACAGCCCATCCATGCATCA	CCCTCACAAGACCTTGGGTATGCTTGA	CC 2104	

Db	33382	TGCTGGGGGAAAGCGCATCCAGCATACCTCAAGAGCCCTGGGATGCTGAC	32441
Qy	2105	ATTGAGGGGACGAGAAAGGTGTCTCTGGAGACTGGTGGGTCTTATGCTTACTTTC	2164
Db	32442	ATTGAGGGGACAGG - AAGTTGTCTCTGGAGACTGGTGGGTCTTATGCTTACTTTC	32500
Qy	2165	TGTCCCGGAACTGTCCTCCAGATCTGTCACTATTCTGAGGGGGTCCNTAAGCGGCA	2224
Db	32501	TGTCCCGGAACTGTCCTCCAGATCTGTCACTA - TCTGAGGGGGTCC - TAAAGCGGCA	32558
Qy	2225	GTCACTAGATCTCTTCCAGGCACTAATTAAGACTGGGGAGCTTATCTTTTAC	2284
Db	32559	GTCACTAGATCTCTTCCAGGCACTAATTAAG - ACTGGGGAGCTTATCTTTTAC	32610
Qy	2285	ATGCTTTTAAATTAATGCTTGAAGGCCCACTACTCTTGTAGGAGAGACATTCAGCA	2344
Db	32617	ATGCTTTTAAATTAATGCTTGAAGGCCCACTACTCTTGTAGGAGAGACATTCAGCA	32676
Qy	2345	AAGCAGGGGCATTATATACCTGAACATAGAGAGAGACACCCGTTGTGTGCCCTG	2404
Db	32677	AAGCAGGGGCATTATATACCTGAACATAGAGAGAGACACCCGTTGTGTGT - CCCCCTG	32739
Qy	2405	CTTGAGGAGGAAATTAATCTGTAAGTCTGGGGCAACAGAGCAATATGACAGGCCAA	2464
Db	32736	CTTGAGGAGGAAATTAATCTGTAAGTCTGGGGCAACAGAGCAATATGACAGG - CAAA	32794
Qy	2465	GATGCGCGTCTCGTTCAAGTTAACTAAAGATTCCACTTCCCTCAACAAAGCA	2524
Db	32795	GATGCGCGTCTCGTTCAAGTTAACTAAAGATTCCACTTCCCTCAACAAAGCA	32854
Qy	2525	GTACCCCTCAAGCCCAAGGCCCAACAGAGATTCAAAAGATGTTAAAGACTTAAAGC	2584
Db	32855	GTACCCCTCAAGCCCAAGGCCCAACAGAGATTCAAAAGATGTTAAAGACTTAAAGC	32914
Qy	2585	CGAAGCTTAGTAAACCATGCACTAATCCCTCGAGTAAATTCGTAGTGAATGAGAG	2644
Db	32915	CGAAGCTTAGTAAACCATGCACTAATCCCTCGAGTAAATTCGTAGTGAATGAGAG	32968
Qy	2645	CACAGAAACCCAGTGGACAGTGGAGGTTAGTCAAGATTCAGAGATTATCATGAGAGC	2704
Db	32968	TACAGAAACCCAGTGGACAGTGGAGGTTAGTCAAGATTCAGAGATTATCAT - GAGGC	33025
Qy	2705	CGTTGTCTTTTATACCAAGCTGTACCTTACCTTATCTGTGTTCCTTCCAAATACGA	2764
Db	33026	TGTTGTCTTTTATACCAAGCTGTACCTTACCTTATCTGTGTTCCTTCCAAATACGA	33085
Qy	2765	GGAAGAGAGTGTTTACASTCTGTGACCTTMAAGAGCTTCTTGTGCATCCCTGTACA	2824
Db	33086	GGAAGAGAGTGTTTACAGTCTGTGACCTTMAAGAGCTTCTTGTGCATCCCTGTACA	33145
Qy	2825	TCTGTACTCAATCTTGTGTTCCTTGAAGTACTTCAAAACCCATCTCAACTCAC	2884
Db	33146	TCTGTACTCAATCTTGTGTTCCTTGAAGTACTTCAAAACCCATCTCAACTCAC	33205
Qy	2885	GTGACATTTTATCCCAAGGTTCAAGGATAGCCCATCTATTTGGCCAGGCAATTAGC	2944
Db	33206	GTGACATTTTATCCCAAGGTTCAAGGATAGCTCCCATCTATTTGGCCAGGCAATTAGC	33265
Qy	2945	CCAAAGCTTAGATCAATTTTCACTGTGACATCTTGTCTCTGTGTCTGTGTGTGATTT	3004
Db	33266	CCAAAGCTTAGATCAATCTTCACTGTGACATCTTGTCTCTGTGTCTGTGTGTGATTT	33323
Qy	3005	ACTTTTGGCGCCYRTTACAAAACCTTGTGCATCAAGCCCAAGCAGTCTTAAATTT	3064
Db	33324	ACTTTTGGCGCCCATTTACAAAACCTTGTGCATCAAGCCCAAGCAGTCTTCAATTT	33383
Qy	3065	CCTCGATCTGTGGCTACAGGTTTCAAGSARABGCTCACTGTCTCAAGCAGGT	3124
Db	33384	CCTCGATCTGTGGCTACATGGTTTCAAAACCAAAGGCTCAACTGTCTCAAGCAGGT	33443
Qy	3125	TAAATACCTTAGGCTTAATTTATCCAAAGCACAGGCCCTCAGTGAAGAAATTAATCA	3184



Db 33444 ----TACTTAGGGCTAAATTTATCCAAAGCCACAGGCGCTCAGTGAAGAACATCCCA 33499  
Qy 3185 GCGTATAGTGGCTTATCCTCATCYCAAAACCTTAAGCACTAAGRGRTTCTTGGCCT 3244  
Db 33500 GCGTATAGTGGCTTATCCTCATCCCAAAACCTTAAGCACTAAGGGGATTCCTTGGCGT 33559  
Qy 3245 AAYAGGTTCTGCCGAAMATGATTCGCCAGTGTGCGAAATATGCCAGTCTTAMATA 3304  
Db 33560 AATAGGTTCTGCCGAAMATGATTCGCCAGTGTGCGAAATATGCCAGTCTTAMATA 33618  
Qy 3305 CAAATAATTAAGAACTCAGAAAGCCATCCATTAATTAAGTGAATGATGAAAGTGAAGTGA 3364  
Db 33619 CAAATAATTAAGAACTCAGAAAGCCATCCATTAATTAAGTGAATGATGAAAGTGAAGTGA 33678  
Qy 3365 AAGTGCTTTCCAGCGCCCTTAAGAAAGGCTTAACCCCAAGCCAGTGTAAAGTTGGCC 3424  
Db 33679 AAGTGCTTTCCAGCGCC-----TAAACCAAGCCAGTGTAAAGTTGGCC 33724  
Qy 3425 AACGGGCAAGACTTTTSTTATATRTCAAGAAAAAACAAGAAATAGCTTGGAGTCC 3484  
Db 33725 AACAGGCAAGACTTTTCTTATATGTCACAG-AAAAAACAAGAAATAGCTTGGAGTCC 33783  
Qy 3485 TTAACAGRTCCAGGGAYAGGCTTGCAACCTATGCGTGAAGTGAAGTGAAGTGAAGTGA 3544  
Db 33784 TTAACAGRTCCAGGGATGAGCTTGCAACCTATGCGTGAAGTGAAGTGAAGTGAAGTGA 33843  
Qy 3545 TAGTGCAAAAGGGTGTGCTCATTTGTTTAYGGGTAGTGTGCAAGTGAAGTGAAGTGAAGTGA 3504  
Db 33844 TAGTGCAAAAGGGTGTGCTCATTTGTTTAYGGGTAGTGTGCAAGTGAAGTGAAGTGAAGTGA 33903  
Qy 3605 CTGAAGCAAGTAAATATATACAGGAGAGATCTTACTGTGTGCACTCTCAAGATGA 3664  
Db 33904 CTGAAGCAAGTAAATATATACAGGAGAGATCTTACTGTGTGCACTCTCAAGATGA 33963  
Qy 3665 AYRGCTACTCACGTCTTAAGAGAGACTGTGGCTGTGAGCAACCTGTCTTAATATTC 3724  
Db 33964 AYRGCTACTCACGTCTTAAGAGAGACTGTGGCTGTGAGCAACCTGTCTTAATATTC 34023  
Qy 3725 AGGCTCTATTACTTGAAGGCGCAGTGTGCTGACCTGTGCACTTGTCACTTTAACCCAG 3784  
Db 34024 AGGCTCTATTACTTGAAGGCGCAGTGTGCTGACCTGTGCACTTGTGCACTTTAACCCAG 34083  
Qy 3785 YNCATTTTCTTCCAGCAATAGAAAGAAAGATATATATCTGTCAACATTAATTTCTC 3844  
Db 34084 CCAATTTTCTTCCAGCAATAGAAAGAAAGATATATATCTGTCAACATTAATTTCTC 34143  
Qy 3845 AAACCTATGCGCACTCGAGGGGACCTTGTAGAGTTCCTTGAAGTGTGCACTTGTCAAC 3904  
Db 34144 AAACCTATGCGCACTCGAGGGGACCTTGTAGAGTTCCTTGAAGTGTGCACTTGTCAAC 34202  
Qy 3905 TTGTATAGTGAAGTTCCTTTGTAGAAAGAAAGAGCTTGAAGAGTGTGAGTGAAGTGA 3964  
Db 34203 TTGTATAGTGAAGTTCCTTTGTAGAAAGAAAGAGCTTGAAGAGTGTGAGTGAAGTGA 34262  
Qy 3965 GTCAAGTATAGGAATATTTGAAGTATCCCTCACTCGAGAACTAAGTGTGAGTGA 4024  
Db 34263 GTCAAGTATAGGAATATTTGAAGTATCCCTCACTCGAGAACTAAGTGTGAGTGA 34322  
Qy 4025 GCGAAGCTAATAGGCTCAATKGGGCACTGAATTTAGAGAAAGAAAGGAGTGAATATA 4084  
Db 34323 GCGAAGCTAATAGGCTCACTTGCGCACTGAATTTAGAGAAAGAAAGGAGTGAATATA 34382  
Qy 4085 TATACAGACTCTATATAGTCTATGCTAGTCTTCATGCGCACTGAGCAATATGAGAGAA 4144  
Db 34383 TATACAGACTCTATATAGTCTATGCTAGTCTTCATGCGCACTGAGCAATATGAGAGAA 34442  
Qy 4145 AGGGAATTTCTTAATCTTGAAGAGAACTATCAACATCAAGAAAGCTTATGAGATTA 4204  
Db 34443 AGGGAATTTCTTAATCTTGAAGAGAACTATCAACATCAAGAAAGCTTATGAGATTA 34502  
Qy 4205 TTATGTGCTGACGAAGAACTATAGAGTGTGAGTGAAGTGTGAGTGTGAGTGTGAGTGTG 4264  
Db 34503 TTATGTGCTGACGAAGAACTATAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 34562

Qy 4265 GGAAGGAAGAAAGGAAATATAGAGAAATGCGCAAGCAKATTTGAAGCAAAAGAGTGA 4324  
Db 34563 GGAAGGAAGAAAGGAAATATAGAGAAATGCGCAAGCAAGATTTGAAGCAAAAGAGTGA 34622  
Qy 4325 AGGAGGAGCCCTCAATTAAGAAATGCTTATTAACCTTCCCTAGTATAGGATTAATCCCTTC 4384  
Db 34623 AGGAGGAGCCCTCAATTAAGAAATGCTTATTAAGAAATGCTTATTAAGGATTAATCCCTTC 34682  
Qy 4385 CGGGAAGCAAGCCCAAGTACTCAGAGAGAAACAGAAATGGGAACTTCAAGAG- CAG 4443  
Db 34683 CGGGAAGCAAGCCCAAGTACTCAGAGAGAAACAGAAATGGGAACTTCAAGAGAG- CAG 34742  
Qy 4444 TTTTCTCCCTCGGAGCGGTTGCGCACTGAAGAGGAAATATCTTTGCTGCACTAT 4503  
Db 34743 TTTTCTCCCTCGGAGCGGTTGCGCACTGAAGAGGAAATATCTTTGCTGCACTAT 34802  
Qy 4504 CCAATGGAATTTACTTAACCTTCACTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 4563  
Db 34803 CCAATGGAATTTACTTAACCTTCACTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 34862  
Qy 4564 TCATATGCGCAATCATTTATTAATGAGCAAGGCTTTTCAAACTATCAAGCAATAT 4623  
Db 34863 TCATATGCGCAATCATTTATTAATGAGCAAGGCTTTTCAAACTATCAAGCAATAT 34922  
Qy 4624 CAGGCGCTGTGAATGTGCGAABAAATATCCCTGCTATATGCGCAAGCTCTTCAGG 4683  
Db 34923 CAGGCGCTGTGAATGTGCGAABAAATATCCCTGCTATATGCGCAAGCTCTTCAGG 34982  
Qy 4684 ARAACAAABAACAGGCTATTAACCTGABABAACCTGCAACCTGATTTTACCAAGCC 4743  
Db 34983 ARAACAAABAACAGGCTATTAACCTGABABAACCTGCAACCTGATTTTACCAAGCC 35042  
Qy 4744 AAACCTCAGGAATTCAGATATCTATGCTGTGCTATATCTTCAAGGTTGGGCAAG 4803  
Db 35043 AAACCTCAGGAATTCAGATATCTATGCTGTGCTATATCTTCAAGGTTGGGCAAG 35102  
Qy 4804 GCGTCCCTGTGAGGACAGAAAGGCGCAAGAGTATTAAGGCACTAGTTCAATGAATA 4863  
Db 35103 GCGTCCCTGTGAGGACAGAAAGGCGCAAGAGTATTAAGGCACTAGTTCAATGAATA 35162  
Qy 4864 ATTCCAGATTCGACTTCCCGAGGCTTACAGATGAACATAGCCCTGCTTTCAGGCC 4923  
Db 35163 ATTCCAGATTCGACTTCCCGAGGCTTACAGATGAACATAGCCCTGCTTTCAGGCC 35222  
Qy 4924 ACAATTAACCAAGGAGTATCCAGGCTTATGATATGATATCACTTCACTGCGCTGA 4983  
Db 35223 ACAATTAACCAAGGAGTATCCAGGCTTATGATATGATATCACTTCACTGCGCTGA 35282  
Qy 4984 AGGCCAAGTCTCAGGAAAGGTCGAAAGAAATGAAGAAATGCTCAAAAGGACATCTTAA 5043  
Db 35283 AGGCCAAGTCTCAGGAAAGGTCGAAAGAAATGAAGAAATGCTCAAAAGGACATCTTAA 35342  
Qy 5044 AAGCAAAACCAAGGAAACCACTCACTGAGTGTGCTGTGCTATAGCTTAAAGAA 5103  
Db 35343 AAGCAAAACCAAGGAAACCACTCACTGAGTGTGCTGTGCTATAGCTTAAAGAA 35402  
Qy 5104 ATCTGCACTTTCCTCAAAAGAGAGACTTGAAGCCATCAAGATGCTGTATGAGAGGCC 5163  
Db 35403 ATCTGCACTTTCCTCAAAAGAGAGACTTGAAGCCATCAAGATGCTGTATGAGAGGCC 35462  
Qy 5164 TTCTATACCAATGACTTGTGCTTGAAGCCAGAGCAAGCTTATAGTGAAGACATCACT 5223  
Db 35463 TTCTATACCAATGACTTGTGCTTGAAGCCAGAGCAAGCTTATAGTGAAGACATCACT 35522  
Qy 5224 CTTAGCAAAATATCAAGAGTCTTAAAGCAATTAACAGAGACCTATCCCTGAGAGAGG 5283  
Db 35523 CTTAGCAAAATATCAAGAGTCTTAAAGCAATTAACAGAGACCTATCCCTGAGAGAGG 35582  
Qy 5284 GAAAGAACTATTCACCCGAGTGAAGTGTATTAAGTGAAGTGTGCTTCTCTATATTC 5343  
Db 35583 GAAAGAACTATTCACCCGAGTGAAGTGTATTAAGTGAAGTGTGCTTCTCTATATTC 35642

OY	5344	CATCCCGAGATACATCCCGGGGAAGGACCCCTACCCGATATTTATATATCCCAACCTGGG	5403
Db	35643	CATCCCTAGATACATCCCGGAAGGACCTTACCCGATATTTATCTACCCCACTGGCG	357020
OY	5404	TTAAAGTGAGTGAGTGAGTCTTGGAATACATCACTGTGAGTCAAAATCCTGATATGCG	5463
Db	35703	TTAAAGTGAGTGAGTGAGTCTTGGAATACATCACTGTGAGTCAAAATCCTGATATGCG	35762
OY	5464	CAAAAGAACTGAAAAATCCAGAGACAAAGCTAGTATTCCTGGAACCTCTAGAGATT	5523
Db	35763	CAAAAGAACTGAAAAATCCAGAGACAAAGCTAGTATTCCTGGAACCTCTAGAGATT	358222
OY	5524	TGAGCGCGCTCTTCAAAACAACACGAGGAAAGTAACTAAATCATTAATCCCCCATG	5583
Db	35823	TGAGCGCGCTCTTCAAAACAACACGAGGAAAGTAACTAAATCATTAATCCCCCATG	35881
OY	5584	GSCTCCCTATCATATTTTTCTCTKATSTGTTTAACTCTSTTCACTCTCATGCA	5643
Db	35882	GCCTCCCTATCATATTTTTCTCTTACCTTTTACCTCTTACCTCTTACCTCTGCA	359431
OY	5644	CCCCCTCATGCGCGTGATGACAGTAGCTCCCTTACMAAGTTTCTATGAGAAATG	5703
Db	35942	CCCCCTCATGCGCGTGATGACAGTAGCTCCCTTACMAAGTTTCTATGAGAAATG	36001
OY	5704	CAGCGTCCCGGGAATAATGATGCCCATCGATAGAGTCTTTSIAGGAAACCCCAACC	5763
Db	36002	CAGCGTCCCGGGAATAATGATGCCCATCGATAGAGTCTTTSIAGGAAACCCCAACC	36061
OY	5764	TTCACTGCCACAACCCATATGCCGGAACGTGATCATCTGCGACCTTTTGATGAT	5823
Db	36062	TTCACTGCCACAACCCATATGCCGGAACGTGATCATCTGCGACCTTTTGATGAT	36121
OY	5824	GCAATATCTCATTATTGAGACAGAAAAATGAATTAATCCTAGTTGTCTGAGAGACTTGGA	5883
Db	36122	GCAATATCTCATTATTGAGACAGAAAAATGAATTAATCCTAGTTGTCTGAGAGACTTGGA	36181
OY	5884	GTCACTGTCTGTTGGACTTACTTACCCAAACGTGATGTCTGATGGGGTGGAGTTCAA	5943
Db	36182	GTCACTGTCTGTTGGACTTACTTACCCAAACGTGATGTCTGATGGGGTGGAGTTCAA	36241
OY	5944	GATCAGGCAAGAAAAACATGTAAAAAGATATCTCCCAACATCACGSGGGTACATGGC	6003
Db	36242	GATCAGGCAAGAAAAACATGTAAAAAGATATCTCCCAACATCACGSGGGTACATGGC	36301
OY	6004	ACCTCTAGCCCTTACAAAGAGCTAGTCTCTCAAACTCATCTGAGAAACCCCTCGTACCAT	6063
Db	36302	ACCTCTAGCCCTTACAAAGAGAGCTAGTCTCTCAAACTCATCTGAGAAACCCCTCGTACCAT	36361
OY	6064	ACTCGCTGGTAAGCTATTTAATACACCTCACTGGGCTCATAGAGTCTCGGCCAA	6123
Db	36362	ACTCGCTGGTAAGCTATTTAATACACCTCACTGGGCTCATAGAGTCTCGGCCAA	36421
OY	6124	AACCCCTACTACGTGGATATGCTCCCTGAACTTCAAGGCATATGTTTCAATCCCT	6183
Db	36422	AACCCCTACTACGTGGATATGCTCCCTGAACTTCAAGGCATATGTTTCAATCCCT	36481
OY	6184	GTACCTGAACATAGAACCACTTCAGACAGAAATTAACAACATTCGCTTTAGTAGGA	6243
Db	36482	GTACCTGAACATAGAACCACTTCAGACAGAAATTAACAACATTCGCTTTAGTAGGA	36541
OY	6244	CCTCTTGTTCATSTGGAATPACCCATACCTCAAACTCACTGATGTAAAAATTTAGC	6303
Db	36542	CCTCTTGTTCATSTGGAATPACCCATACCTCAAACTCACTGATGTAAAAATTTAGC	36601
OY	6304	AATTACTACATACCAACCACTCCCAATGATCAGGGGAGTACCTCCGCCAACAAATA	6363
Db	36602	AATTACTACATACCAACCACTCCCAATGATCAGGGGAGTACCTCCGCCAACAAATA	36661
OY	6364	GTCTGCTACCCCTCAGGAATATTTTGTCTGTGTACTCAGGCTATCGTTGTTGAT	6423
Db	36662	GTCTGCTACCCCTCAGGAATATTTTGTCTGTGTACTCAGGCTATCGTTGTTGAT	36721
OY	6424	GGCCTTCAGAAATCTATGAGCTTCTCTCATTTCTTAGTCCCCCATATGRCATCTACA	6483

Db	36722	GGCTCTTCAAGAACTAATGCTCTCCCTCATTTCTAAGTACCCCTAATGACCACTACACT	36781
Qy	6484	GAACAGATTATTAACGTTATGTCAATCTAAGACCCCGCAACAAAGATACCATCTT	6543
Db	36782	GAACAAGATTATTAAGTTATGTCAATCTAAGACCCCGCAACAAAGATACCATCTT	36841
Qy	6544	CCTTTGTATTAGAGCAGAGAGTGTAGTGCATAGTACTGGCACTTGGCGTATACA	6603
Db	36842	CCTTTGTATTAGAGCAGAGAGTGTAGTGCATAGTACTGGCACTTGGCGTATACA	36901
Qy	6604	ACCTCACTCACTTCTACAACTAATCTCAAGAACTAAATGGGGACATGGAACGGGTC	6663
Db	36902	ACCTCACTCACTTCTACAACTAATCTCAAGAACTAAATGGGGACATGGAACGGGTC	36961
Qy	6664	GCCGACTCCCTGGTCACTTGGCAATCACTTAACTCCCTAGACAGACTAGTCTTCA	6723
Db	36962	GCCGACTCCCTGGTCACTTGGCAATCACTTAACTCCCTAGACAGACTAGTCTTCA	37021
Qy	6724	AATCGAAGACTTATAGACTTGTAAACCGCTGAPAGAGGGGGAACTGTTATTTTGGG	6783
Db	37022	AATCGAAGACTTATAGACTTGTAAACCGCTGAPAGAGGGGGAACTGTTATTTTGGG	37081
Qy	6784	GAAGAATGTGTATTATGTATTAATCAATCCGGAATGTCACTGAGAAATTTAAGAAAT	6843
Db	37082	GAAGAATGTGTATTATGTATTAATCAATCCGGAATGTCACTGAGAAATTTAAGAAAT	37141
Qy	6844	CSAGATCGAATCAACGTAKAGCAGABAGCTTCCAAAACATGACACCCCTGGGCGTCTC	6903
Db	37142	CGAGATCGAATCAACGTAKAGCAGABAGCTTCCAAAACATGACACCCCTGGGCGTCTC	37201
Qy	6904	AGCCATGAGATGCCCTGATATTCCTCCCTTCTTGAAGACCTTGAAGACTAATATTTGCTA	6963
Db	37202	AGCCATGAGATGCCCTGATATTCCTCCCTTCTTGAAGACCTTGAAGACTAATATTTGCTA	37261
Qy	6964	CTCCCTTTTGGACCCCTGTATCTTTAACCCTCTTGTATACTTGTCTCTCCAGATTCGA	7023
Db	37262	CTCCCTTTTGGACCCCTGTATCTTTAACCCTCTTGTATACTTGTCTCTCCAGATTCGA	37321
Qy	7024	GCTGTAAACTCAAAATGAGAGCCCAAGATGACCTCAAGACTTAAGCTTACCGCAGACCC	7083
Db	37322	GCTGTAAACTCAAAATGAGAGCCCAAGATGACCTCAAGACTTAAGCTTACCGCAGACCC	37381
Qy	7084	CTGACCGGCGCTGTAGGCCCAAGATCTGATGTTAATGAATCACTAAAGGCAACCCCTCTGAG	7143
Db	37382	CTGACCGGCGCTGTAGGCCCAAGATCTGATGTTAATGAATCACTAAAGGCAACCCCTCTGAG	37441
Qy	7144	GAAATCTCAGCTGCACAACTCTTACTAAGCCCCAAATTGACGAGGAAGCTTAAAGCCGT	7203
Db	37442	GAAATCTCAGCTGCACAACTCTTACTAAGCCCCAAATTGACGAGGAAGCTTAAAGCCGT	37501
Qy	7204	SGTGGCCCAACCTCCCAACAGACTTAAGTTTTCTGTGAGATGGGGGACCTGAGAGAC	7263
Db	37502	SGTGGCCCAACCTCCCAACAGACTTAAGTTTTCTGTGAGATGGGGGACCTGAGAGAC	37561
Qy	7264	AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCTTAGSTGGGAAGGTGACC	7323
Db	37562	AGGACTAGCTGATTTCTTAGCTGATTAAGAAATCCYTAAGCTTAGSTGGGAAGGTGACC	37621
Qy	7324	ACATCACTTTTAAACATGGGGCTTGCACCTTAATGTTCAACCTGACCAATCAGAGAGCTC	7383
Db	37622	ACATCACTTTTAAACATGGGGCTTGCACCTTAATGTTCAACCTGACCAATCAGAGAGCTC	37681
Qy	7384	ACTAAATAGCTAATTAGGCAAAAGACGAGAGGTAAGAAATAGCAATCATATTAATGCMTG	7443
Db	37682	ACTAAATAGCTAATTAGGCAAAAGACGAGAGGTAAGAAATAGCAATCATATTAATGCMTG	37741
Qy	7444	AGAGCAGCAGAGAGGCAATGATCGGGAATATAAACCAAGATTTTGAAGCCGGCAACGG	7503
Db	37742	AGAGCAGCAGAGAGGCAATGATCGGGAATATAAACCAAGATTTTGAAGCCGGCAACGG	37801
Qy	7504	CAACCCCTTTGGGTCCCTCCCTTGTATGAGAGCTCTGTTTTCAATGCTATTTCACTCT	7563

Db 37802 CAACCCCTTGCGTCCCTCCCTTGATGAGAGCTCTGTTTCATGCTATTCTACTCT 37861  
 Oy 7564 ATTAATCTTGACATGCR 7582  
 Db 37862 ATTAATCTTGCAACTGCA 37880

RESULT 9  
 AC007566/c 149194 bp DNA linear PRI 01-MAR-2002  
 LOCUS Homo sapiens BAC clone CTB-10G5 from 7q21-7q22, complete sequence.  
 AC007566  
 AC007566.2 GI:11181861  
 VERSION AC007566.2  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Suleston,J.E. and Waterston,R.  
 TITLE Toward a complete human genome sequence  
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
 MEDLINE 99063792  
 PUBMED 9847074

REFERENCE  
 AUTHORS Du,Z.  
 TITLE The sequence of Homo sapiens BAC clone CTB-10G5  
 JOURNAL Unpublished (2001)  
 REFERENCE 3 (bases 1 to 149194)  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-MAY-1999) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA

REFERENCE  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-OCT-2000) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 5 (bases 1 to 149194)  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2000) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA

REFERENCE  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JUN-2002) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 REFERENCE 7 (bases 1 to 149194)  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-FEB-2002) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA

REFERENCE  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 8 (bases 1 to 149194)  
 AUTHORS Waterston,R.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT  
 On Nov 16, 2000 this sequence version replaced g1:4835815.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@wustl.wustl.edu](mailto:sapiens@wustl.wustl.edu)  
 ----- Summary Statistics  
 Center project name: H\_RG010G05

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.

MAPPING INFORMATION:  
 The sequence of this clone was established as part of a mapping and  
 sequencing collaboration between the NHGRI Chromosome 7 Mapping  
 Project (Eric D. Green, Director), John D. McPherson in the  
 Department of Genetics (Washington University), and the Washington  
 University Genome Sequencing Center. For additional information  
 about the map position of this sequence, see  
<http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send  
<mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
 Clone CTB-10G5 is from the first release of the human BAC library  
 CTB-978SK-B. The library contains cloned DNA from the male  
 fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad.  
 Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8  
 (1996). This clone is available from Research Genetics, Inc.  
 (<http://www.reegen.com>).  
 VECTOR: pBelOBAC11  
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the left is RP5-91H5, 200 base pair  
 overlap. Actual start of this clone is at base position 195 of  
 CTB-10G5; actual end is at base position 150532 of CTB-10G5.

The clone CTB-10G5 contains the entire sequence of CTB-83M5.

#### FEATURES

##### source

1..149194  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="7"  
 /map="7q21-7q22"  
 /clone="CTB-10G5"  
 /clone\_id="CTB-978SK-B"

##### repeat\_region

1..1634  
 /rpt\_family="L1"

##### misc\_feature

2248..2388  
 /note="match to EST BG752883 (NID:g14063536)"

##### misc\_feature

2248..2387  
 /note="match to EST AA149693 (NID:g1720635) zn99d06.r1"

##### misc\_feature

2248..2387  
 /note="match to EST AW579261 (NID:g7254310)"

##### misc\_feature

2248..2387  
 /note="match to EST BG766882 (NID:g14077535)"

##### misc\_feature

2248..2387  
 /note="match to EST B1160365 (NID:g14620366)"

##### misc\_feature

2248..2374  
 /note="match to EST B1160365 (NID:g14620366)"

##### misc\_feature

2248..2287  
 /note="similar to Homo sapiens EST BF758865  
 (NID:g12106765)"

##### misc\_feature

2248..2287  
 /note="match to EST AV686676 (NID:g10288539)"

##### misc\_feature

2253..2387  
 /note="match to EST AU123510 (NID:g10948226)"

##### repeat\_region

2344..2387  
 /note="match to EST AV686676 (NID:g10288539)"

##### repeat\_region

2696..3066  
 /rpt\_family="MAIR"

##### misc\_feature

3108..3392  
 /rpt\_family="Alu"  
 3540..3628

misc_feature	/note="match to EST AA425526 (NID: g2106267) zw48D03.r1"
	3540. .3628
	/note="similar to Mus musculus EST BB253526 (NID: g8946272) "
	3542. .3632
misc_feature	/note="match to EST AV686676 (NID: g10208939) "
	3542. .3628
misc_feature	/note="match to EST AA149693 (NID: g1720635) zn99d06.r1"
	3542. .3628
misc_feature	/note="match to EST AU123510 (NID: g10948226) "
	3542. .3628
misc_feature	/note="match to EST BG766882 (NID: g14077535) "
	3542. .3628
misc_feature	/note="match to EST B1160365 (NID: g14620366) "
	3542. .3609
misc_feature	/note="match to EST AW579261 (NID: g7254310) "
	3562. .3628
misc_feature	/note="match to EST BE272564 (NID: g9146913) "
	3598. .3628
misc_feature	/note="match to EST BE299708 (NID: g9183456) "
	3714. .3785
misc_feature	/note="match to EST BG260659 (NID: g12770475) "
	3717. .3785
	/note="similar to Mus musculus EST BE949436 (NID: g10678674) "
	3719. .3785
misc_feature	/note="match to EST AA149693 (NID: g1720635) zn99d06.r1"
	3719. .3785
misc_feature	/note="match to EST AA425526 (NID: g2106267) zw48D03.r1"
	3719. .3785
misc_feature	/note="match to EST AU123510 (NID: g10948226) "
	3719. .3785
misc_feature	/note="match to EST BE272564 (NID: g9146913) "
	3719. .3785
misc_feature	/note="match to EST BE299708 (NID: g9183456) "
	3719. .3785
misc_feature	/note="match to EST BG766882 (NID: g14077535) "
	3719. .3785
misc_feature	/note="match to EST B1160365 (NID: g14620366) "
	3719. .3785
	/note="similar to Mus musculus EST BB253526 (NID: g8946272) "
	3860. .3862
misc_feature	/note="match to EST AV686676 (NID: g10288539) "
	4730. .4827
repeat_region	/fpd_family="L2"
	4834. .4925
repeat_region	/fpd_family="L1"
	5715. .5860
misc_feature	/note="match to EST AA425526 (NID: g2106267) zw48D03.r1"
	5715. .5860
misc_feature	/note="match to EST AU123510 (NID: g10948226) "
	5715. .5860
misc_feature	/note="match to EST BE272564 (NID: g9146913) "
	5715. .5860
misc_feature	/note="match to EST BE299708 (NID: g9183456) "
	5715. .5860
misc_feature	/note="match to EST BG260659 (NID: g12770475) "
	5715. .5860
misc_feature	/note="match to EST BG766882 (NID: g14077535) "
	5715. .5860
misc_feature	/note="match to EST B1160365 (NID: g14620366) "
	5715. .5860

	Query Match	Similarity	85.0%	Score 6448	DB 9	Length 149194
Beet	Local	Conservatively	95.7%	Pred.	No. 0	
Matches 6681	Conservative	195	Mismatches	50	Indels	53
					Gaps	27
Oy	606	GATGGGAAACGTTCCCGCGAAGACAAAACGCCCTTAAGACGTATTCTGARAATTGGCA	665			
Db	91051	GATGGGAAACGTTCCCGCGAAGACAAAACGCCCTTAAGACGTATTCTGAGAATTGGGA	90929			
Oy	666	MCAATTGACCGCTCAGACACTPAAGAAAGAAAGCACTTATATTCTTCGACAGTCCGCGCT	725			

Db	90991	CGAATTGACCTTCAGACACTTAAGAAAGAAAGCACTTAATATCTTTGAGTGCGCCCTG	90932
Qy	726	GCACCTCCGAGGGGAAGTATAAATTAACACCATCTTACAGGTAGACYTCTTTGTAGAA	785
Db	90991	GCACTCTGAGGGGAAGTAAATTAATTAACACCATCTTACAGGTAGACCTTTTGTAGAA	90872
Qy	786	AAGGCAAAATGAGTGAAGTGCCATTAAGTCAAACTTTCTTTTCATTTAAGAGCAACTAC	845
Db	90871	AAGGCAAAATGAGTGAAGTGCCATTAAGTCAAACTTTCTTTTCATTTAAGAGCAACTAC	90812
Qy	846	AATTATGTAAAAAGTGGATTTATGCCCCATACAGAAAGCCTTCAGAGTCTAAGCTCCCTATC	905
Db	90811	AATTATGTAAAAAGTGGATTTATGCCCCATACAGAAAGCCTTCAGAGTCTAAGCTCCCTATC	90755
Qy	906	CCAGCATCCCCGACCTCCCTCCCAATTAATTAAGAACCCCTTCACAAACCAATGGTCCAA	965
Db	90751	CCAGCATCCCCGACCTCCCTCCCAATTAATTAAGAACCCCTTCACAAACCAATGGTCCAA	90699
Qy	966	AAGGAGATGACAAAAAGGTTAAACAGTGAACCAAGAGTGCCATATTTCCCAATTATGA	1025
Db	90691	AAGGAGATGACAAAAAGGTTAAACAGTGAACCAAGAGTGCCATATTTCCCAATTATGA	90632
Qy	1026	CCCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCAGCAGAGTGCATGTCCTTTTAY	1085
Db	90631	CCCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGCCAGCAGAGTGCATGTCCTTTTTC	90573
Qy	1086	TCTCCGACCTTAAGCAAAATTAATAACAGACTTGAATTAATTCAGATAATCTGATAG	1145
Db	90572	TCTCCGACCTTAATTAAGCAAAATTAATAACAGACTTGAATTAATTCAGATAATCTGATAG	90513
Qy	1146	CTATATATGATGTTTTCAAAGGTTAGGACAAATTCCTTGATCTGACATGGAGAGATATATA	1205
Db	90512	CTATATATGATGTTTTCAAAGGTTAGGACAAATTCCTTGATCTGACATGGAGAGATATA-A	90454
Qy	1206	TGTCACTGTCTAAATTCAGACACTTAACCCCAATAGAGAAAGTGCACATTAATCTGCAGCT	1265
Db	90453	TGTCACTGTCTAAATTCAGACACTTAACCCCAATAGAGAAAGTGCACATTAATCTGCAGCT	90394
Qy	1266	GAGGTGTTGGCGATTCCTGGTATCTCAGTCAGGTCAATGATATANGATATACAACGAAGG	1325
Db	90393	GAGGTGTTGGCGATTCCTGGTATCTCAGTCAGGTCAATGATATANGATATACAACGAAGG	90337
Qy	1326	AAAGANAATATTTCCCAACAGCCAGCAGACAGATTTCCCAAGTCTAATACCTCATTTGGGAC	1385
Db	90336	AAAGANAATATTTCCCAACAGCCAGCAGACAGATTTCCCAAGTCTAATACCTCATTTGGG--	90280
Qy	1386	ACAGAAATCAGTAACATGGGAGATTGGTCTGCAACATTTGCTTAATTCTGTGCTAATA	1445
Db	90279	ACAGAAATCAGTAACATGGGAGATTGGTCTGCAACATTTGCTTAATTCTGTGCTAATA	90221
Qy	1446	GGAATTAAGGAAACTASGAAAGAAATCTATGAAATTAATCTAATGATGTCCACATTAACACA	1505
Db	90220	GGAATTAAGGAAACTASGAAAG--AAGTCTATGAAATTAATCTAATGATGTCCACATTAACACA	90162
Qy	1506	GGGGAAGGGAAGAAATTCCTACTGCTTTCTGGAGAGACTAAGGGAGGCAATTGAGGAAGC	1565
Db	90161	GGGAAGGGAAGAAATTCCTACTGCTTTCTGGAGAGACTAAGGGAGGCAATTGAGGAAGC	90103
Qy	1566	GTCGCTCTCTGTCACTGACTCTTCTGAAGGCCACTTAATCTTAAAGGTAAAGTTTATCA	1625
Db	90102	GTCGCTCTCTGTCACTGACTCTTCTGAAGGCCACTTAATCTTAAAGGTAAAGTTTATCA	90043
Qy	1626	CTCAGTAGCTGACGACATTAG--AAAAAATCTCAAAAGTCTGCGTAGGCGCGAGCAAA	1684
Db	90042	CTCAGTAGCTGACGACATTAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	89983
Qy	1685	ACTTAGAAACCTTATTTGAATTTGGCACTTGGGTTTTTATATATAGAGTCAAGAGGAGC	1744
Db	89982	ACTTAGAAACCTTATTTGAATTTGGCACTTGGGTTTTTATATATAGAGTCAAGAGGAGC	89923
Qy	1745	AGCGGAACGAGCAAAAGGATTAATAAAAGGCCACGGCTTATGCTATGACCTTCAGG	1804

Dh 89922 AGCGGAAACAGACAAACGGGATTTAAAAAAGGCCACCGCTTATGATGACCTTCAGG 89863  
Qy 1805 CAAGTGAATTTGGAGGCTCTGGAAAAAGGAAAAAGTGGCAATTAATGATGCTTAATAG 1864  
Db 89862 CAAGTGAATTTGGAGGCTCTGGAAAAAGGAAAAAGTGGCAATTAATGATGCTTAATAG 89803  
Qy 1865 GCTTGCTTCCAGTGGCGGTCTACAGGACATTTAAAAAAGATTGTCCAGTAAAGTAA 1924  
Db 89802 GCTTGCTTCCAGTGGCGGTCTACAGGACATTTAAAAAAGATTGTCCAGTAAAGTAA 89743  
Qy 1925 CGGCCCCCTGTGTCAATGCCCCCTTAATTTCAAGGGAATACATGGAAGGCCCATGCCCCAG 1984  
Db 89742 CGGCCCCCTGTGTCAATGCCCCCTTAATTTCAAGGGAATACATGGAAGGCCCATGCCCCAG 89683  
Qy 1985 GGACAAAGGCTTTTATGAGTCAAGGCACTAACCATGATCCAGACGACGAGCTAGAG 2044  
Db 89682 GGACAAAGGCTTTTATGAGTCAAGGCACTAACCATGATCCAGACGACGAGCTAGAG 89623  
Qy 2045 TGCCTGGGGCAAGCGCCCATCCATGCGCATCACTCTACAGAGCCCTGGGTATGCTGACC 2104  
Db 89622 TGCCTGGGGCAAGCGCCCATCCATGCGCATCACTCTACAGAGCCCTGGGTATGCTGACC 89563  
Qy 2105 ATTTAGAGGCGCAGAGAGTGTGTCTCCGAGACATGSGTGGCTTTCTTAATCTTACTTC 2164  
Db 89562 ATTTAGAGGCGCAGAGAGTGTGTCTCCGAGACATGSGTGGCTTTCTTAATCTTACTTC 89504  
Qy 2165 TGTCCCGGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGTCTCCTTAAGACGGGCA 2224  
Db 89503 TGTCCCGGACAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGTCTCCTTAAGACGGGCA 89446  
Qy 2225 GTTCACTAGATCTTTTATCCAGCACTAAGTTATGAATCTGGGAGCTTTATTTCTTTTAC 2284  
Db 89445 GTTCACTAGATCTTTTATCCAGCACTAAGTTATGAATCTGGGAGCTTTATTTCTTTTAC 89388  
Qy 2285 ATGCTTTTCTAATTAAGTCTGAAAAGGCCCACTACCTTTAGGAGAGACATTTAGCA 2344  
Db 89387 ATGCTTTTCTAATTAAGTCTGAAAAGGCCCACTACCTTTAGGAGAGACATTTAGCA 89328  
Qy 2345 AAGCAGAGGCGCATTTATACCTGTACATAGAGAGAGAAACCGGTTTGTGTGTCCTG 2404  
Db 89327 AAGCAGAGGCGCATTTATACCTGTACATAGAGAGAGAAACCGGTTTGTGTGTCCTG 89269  
Qy 2405 CTTGAGAGAGAAATTAATCTCTGAATGTGGGCAACAGAGACAAATTTGACGAGCCAA 2464  
Db 89268 CTTGAGAGAGAAATTAATCTCTGAATGTGGGCAACAGAGACAAATTTGACGAGCCAA 89210  
Qy 2465 GAATGCGCGTCTGTTCAAGTTAACTAAAGATTCACATTCCTTTCCCTAACCAAGGCA 2524  
Db 89209 GAATGCGCGTCTGTTCAAGTTAACTAAAGATTCACATTCCTTTCCCTAACCAAGGCA 89150  
Qy 2525 GTACCCCTCAGACCCAAAGGCCCAACAGATTCCTAAAGATTTGTTAAGACTTAAGC 2584  
Db 89149 GTACCCCTCAGACCCAAAGGCCCAACAGATTCCTAAAGATTTGTTAAGACTTAAGC 89090  
Qy 2585 CCAAGGCTTATGTAACATGACATTAATCTCTGACATTAATCTCTGATGATGAGAG 2644  
Db 89089 CCAAGGCTTATGTAACATGACATTAATCTCTGACATTAATCTCTGATGATGAGAG 89037  
Qy 2645 CAAAGAAACCAAGTGAAGTGAAGGTTAGTGAAGTCTCAGATTAATCAATGAGG 2704  
Db 89036 CAAAGAAACCAAGTGAAGTGAAGGTTAGTGAAGTCTCAGATTAATCAATGAGG 88979  
Qy 2705 CGTGTCTTTTATACCAAGTGTACTAGCCCTTAATCTGTGTGTTCCCAATACAGA 2764  
Db 88978 TGTGTCTCTCTATAGCAAGTGTACTAGCCCTTAATCTGTGTGTTCCCAATACAGA 88919  
Qy 2765 GGAAGCAGAGTGTATACATCTCTGACATTTAGAGATGCTTTCTTCTGACATCTCTGAC 2824  
Db 88918 GGAAGCAGAGTGTATACATCTCTGACATTTAGAGATGCTTTCTTCTGACATCTCTGAC 88859  
Qy 2825 TCCGATCTCAATCTTGTGTGCTTTGAGATTAATCAATCAACCAATCTCAATCTCA 2884  
Db 88858 TCCGATCTCAATCTTGTGTGCTTTGAGATTAATCAATCAACCAATCTCAATCTCA 88799

Qy 2885 CTGACATTTTATCCCAAGGTTCAAGGATAGTCCCATCTAATTTGCTCAAGCATTTAG 2944  
Db 88798 CTGACATTTTATCCCAAGGTTCAAGGATAGTCCCATCTAATTTGCTCAAGCATTTAG 88739  
Qy 2945 CCAAGCTTGAATCTATTTATCTGAGACCTGTGTCTTCTGTGATGATGAT 3004  
Db 88738 CCAAGCTTGAATCTATTTATCTGAGACCTGTGTCTTCTGTGATGATGAT 88681  
Qy 3005 ACTTTTGGCGCTTTCAGAAACCTGTGCTCATGAGCACCACCAACCTCTTAAATTT 3064  
Db 88680 ACTTTTGGCGCTTTCAGAAACCTGTGCTCATGAGCACCACCAACCTCTTAAATTT 88621  
Qy 3065 CCTCGTACTGTGGCTACATGCTTTCAGAAACCAAGGCTCACTGTCTCAGAGAGT 88561  
Db 88620 CCTCGTACTGTGGCTACATGCTTTCAGAAACCAAGGCTCACTGTCTCAGAGAGT 88504  
Qy 3125 TAAATTAAGTAAATTAATTCAGAAAGGACACAGGCGCTCAGTGAAGAAATATCA 3184  
Db 88560 TAAATTAAGTAAATTAATTCAGAAAGGACACAGGCGCTCAGTGAAGAAATATCA 88505  
Qy 3185 GCTATACTGCTTATCTCATCTCAATCAAAACCTTAAGCACTAAGGATTTCTGCT 3244  
Db 88504 GCTATACTGCTTATCTCATCTCAATCAAAACCTTAAGCACTAAGGATTTCTGCT 88445  
Qy 3245 AATAGGTTTCTGCGAAATGATTTCCAGGTTGGCGAAATAGCAGAGCTTAAATTA 88386  
Db 88444 AATAGGTTTCTGCGAAATGATTTCCAGGTTGGCGAAATAGCAGAGCTTAAATTA 88326  
Qy 3305 CASTAATTAAGAAACTCAGAAAGCCAAATACCATTAATTAAGTGAAGAAATGAAG 3364  
Db 88385 CASTAATTAAGAAACTCAGAAAGCCAAATACCATTAATTAAGTGAAGAAATGAAG 88326  
Qy 3365 AAGTGGCTTTCAGAGCGCCCTAAAGAGGCTTAAACCAAGTCCAGTGTAAATTTG 3424  
Db 88325 AAGTGGCTTTCAGAGCGCCCTAAAGAGGCTTAAACCAAGTCCAGTGTAAATTTG 88280  
Qy 3425 AACGGGCAAGACTTTTSTTATATATCAAGAAAGAAACAGAAATAGCTCTGAGTGC 3484  
Db 88279 AACGGGCAAGACTTTTSTTATATATCAAGAAAGAAACAGAAATAGCTCTGAGTGC 88221  
Qy 3485 TTACACAGTCCBAGGAGAGAGCTTGCACATCTGTGCGATACCTGATGAGAAATGAT 3544  
Db 88220 TTACACAGTCCBAGGAGAGAGCTTGCACATCTGTGCGATACCTGATGAGAAATGAT 88161  
Qy 3545 TAGTGCAAGAGGTTGCTTATTTATGAGTGTGAGTGTGAGTGTGAGTGTGAT 3604  
Db 88160 TAGTGCAAGAGGTTGCTTATTTATGAGTGTGAGTGTGAGTGTGAGTGTGAT 88101  
Qy 3605 CTGACACATTAATTAATTAAGGAGAGTCTTATCTGTGAGACATCTCATGAGTGA 3664  
Db 88100 CTGACACATTAATTAATTAAGGAGAGTCTTATCTGTGAGACATCTCATGAGTGA 88041  
Qy 3665 ATGACATCTCACTGCTTAAGAGACTTGTGTGTGACAGCACTGTTTCTTAATATGTC 3724  
Db 88040 ATGACATCTCACTGCTTAAGAGACTTGTGTGTGACAGCACTGTTTCTTAATATGTC 87981  
Qy 3725 AGGCTCTATTAATTAAGAGGCGAGTGTGCTGACATGTCATCTTAAACCAAG 3784  
Db 87980 AGGCTCTATTAATTAAGAGGCGAGTGTGCTGACATGTCATCTTAAACCAAG 87921  
Qy 3785 YCNCATTTCTTCCAGCAATTAAGAAAGATTAATTAATTAATTAATTAATTTCTC 3844  
Db 87920 YCNCATTTCTTCCAGCAATTAAGAAAGATTAATTAATTAATTAATTTCTC 87861  
Qy 3845 AAACCTATGCACTCAAGGAGACTTATGAGAGTCTTGAAGTGTGAGTGTGAGTGTGAG 3904  
Db 87860 AAACCTATGCACTCAAGGAGACTTATGAGAGTCTTGAAGTGTGAGTGTGAGTGTGAG 87802  
Qy 3905 TTGTATTAATGAAGTGTCTTTTGAAGAAAGACTTGAAGAAAGTGTGTGAGTGTGAG 3964  
Db 87801 TTGTATTAATGAAGTGTCTTTTGAAGAAAGACTTGAAGAAAGTGTGTGAGTGTGAG 87742



```
OY 3965 GTGATGTAATATGGAATATTTGAAAGTAATCCCTCACTCCAGGAACTAAGTCYAGCTR 4024
|||
|||
|||
Db 87741 GTGATGTAATATGGAATATTTGAAAGTAATCCCTCACTCCAGGAACTAAGTCAGCTA 87682
|||
|||
|||
OY 4025 GCAGAACTAATAGACCCYATYTKGSGCACTAGAAATTAGGAGAAAGAAAAAGGYYAATTA 4084
|||
|||
|||
Db 87681 GCAGAACTAATAGACCCCTCACTTGGSCACTAGATTAGAGAAAGAAAAAGGSCAAATATA 87622
|||
|||
|||
OY 4085 TATAACAGACTTATATATGCTYACCTAGTCNTCCATGCGCATGMRGCAATATGSAFAGA 4144
|||
|||
|||
Db 87621 TATACAGACTTAAATATGCTTATCTAGTCCTCCATGCGCAATGAGCAATATAGAAAGA 87562
|||
|||
|||
OY 4145 AGGGAATTCCTTAATCTTCGABGAAACACCTTATCMAACATCAGAAAGCCATTAGARATTA 4204
|||
|||
|||
Db 87561 AGGGAATTCCTTAATCTTCGABGAAACACCTTATCMAACATCAGAAAGCCATTAGAAATTA 87502
|||
|||
|||
OY 4205 TTATYTGCGMTACAGAAACCTTAPAGAGGTGMAAGCTTACACTGCGYGGGTCATCANAAA 4264
|||
|||
|||
Db 87501 TTATYTGCGMTACAGAAACCTTAPAGAGGTGMAAGCTTACACTGCGYGGGTCATCANAAA 87442
|||
|||
|||
OY 4265 GGAAGGAAGAAAGGAAATASAAAGAAATGCCAAGCAKATATTGAAGMAAAAGAGCTGA 4324
|||
|||
|||
Db 87441 GGAAGGAAGAAAGGAAATASAAAGAAATGCCAAGCAKATATTGAAGMAAAAGAGCTGA 87382
|||
|||
|||
OY 4325 AGGCAAGACCTCCATTAAGAAATGCTTATTAACCTTCCCTTAAGTATAGGTTAATCCCTTC 4384
|||
|||
|||
Db 87381 AGGCAAGACCTCCATTAAGAAATGCTTATTAACCAACCCCTAAGTATAGGTTAATCCCTTC 87322
|||
|||
|||
OY 4385 CGGGAAACCAAGCCCGAGTATCTCAGAGAGAAACAGAAATGGGGAACCTCGACGAG- CAG 4443
|||
|||
|||
Db 87321 CGGGAAACCAAGCCCGAGTATCTCAGAGAGAAACAGAAATGGGGAACCTCGACGAG- CAG 87262
|||
|||
|||
OY 4444 TTTTCTCCCTCGGGAAGGTTAGCCACTGAGAAAGAGAAATATCTTTTGCTCGCACTAT 4503
|||
|||
|||
Db 87261 TTTTCTCCCTCGGGAAGGTTAGCCACTGAGAAAGAGAAATATCTTTTGCTCGCACTAT 87202
|||
|||
|||
OY 4504 CCAATGGAATTAATCTTAAACCTTCATCAAACTTTCACTTAGGCATCGATAGACCCA 4563
|||
|||
|||
Db 87201 CCAATGGAATTAATCTTAAACCTTCATCAAACTTTCACTTAGGCATCGATAGACCCA 87142
|||
|||
|||
OY 4564 TCARATGCGCAAAATCATTAATTTACTGAGACGAGCCTTTTCAAACTATCAAGCABATAT 4623
|||
|||
|||
Db 87141 TCARATGCGCAAAATCATTAATTTACTGAGACGAGCCTTTTCAAACTATCAAGCABATAT 87082
|||
|||
|||
OY 4624 CAGGGCTGTGAATGTGCGCARAATAATCCCTGCTATGCGCAAGCTCTTCAAG 4683
|||
|||
|||
Db 87081 CAGGGCTGTGAATGTGCGCARAATAATCCCTGCTATGCGCAAGCTCTTCAAG 87022
|||
|||
|||
OY 4684 ABAACAABAACAGGCCATTACCTGAPABAAPACTGGCAACTGATTTTACCCACAGCCC 4743
|||
|||
|||
Db 87021 ABAACAABAACAGGCCATTACCTGAPABAAPACTGGCAACTGATTTTACCCACAGCCC 86962
|||
|||
|||
OY 4744 AAACCTCAGGGATTTTCAATTAATTTACTGAGACGAGCCTTTTCAAACTATCAAGCABATAT 4803
|||
|||
|||
Db 86961 AAACCTCAGGGATTTTCAATTAATTTACTGAGACGAGCCTTTTCAAACTATCAAGCABATAT 86902
|||
|||
|||
OY 4804 GCCTTCCCTGTAGAGACAGAAAAAGCCCAAGAGTAATAAGGCACTAGTTTCAAGAAATA 4863
|||
|||
|||
Db 86901 GCCTTCCCTGTAGAGACAGAAAAAGCCCAAGAGTAATAAGGCACTAGTTTCAAGAAATA 86842
|||
|||
|||
OY 4864 ATTTCCAGATTCGGAATTCCTCCAGAGCTTACAGAGTACAAATAGCCCTGCTTTCAGAGCC 4923
|||
|||
|||
Db 86841 ATTTCCAGATTCGGAATTCCTCCAGAGCTTACAGAGTACAAATAGCCCTGCTTTCAGAGCC 86782
|||
|||
|||
OY 4924 ACAGTAACCCAGGAGATATCCAGAGGCTTAGGTAATCGATATACATTAACCTCGGCTGA 4983
|||
|||
|||
Db 86781 ACAGTAACCCAGGAGATATCCAGAGGCTTAGGTAATCGATATACATTAACCTCGGCTGA 86722
|||
|||
|||
OY 4984 AGGCCACAGTCTCTCAGGAGAGGTGAGAAATATGATGAATTAAGTCAAGAGCACTCTTAAA 5043
|||
|||
|||
Db 86721 AGGCCACAGTCTCTCAGGAGAGGTGAGAAATATGATGAATTAAGTCAAGAGCACTCTTAAA 86662
|||
|||
|||
OY 5044 AAGCAAACCCAGGAAACCAACCTCACTAGGCTGTCTGTGCTTATAGCTTTAAAAAGA 5103
|||
|||
|||
```

Db 85582 AACCTACTAAGTGGATATGCTCCCTGAACTTCAGGCAATATGTTCAATCCCT 85523  
 Qy 6184 GTAACCTGAACATGGAACAACTTCAGCAGAGAAATAAACACACTTCGTTTATGAGAA 6243  
 Db 85522 GTAACCTGAACATGGAACAACTTCAGCAGAGAAATAAACACACTTCGTTTATGAGAA 85463  
 Qy 6244 CCTCTGTTCATATGGAATTAACCAATACCTCAACCTCAGCTGTGTTAAATTTAGC 6303  
 Db 85462 CCTCTGTTCATATGGAATTAACCAATACCTCAACCTCAGCTGTGTTAAATTTAGC 85403  
 Qy 6304 AATACTACATACACAACTCCCAATGATCAGTGGGATCTCCCTCCACACAATA 6363  
 Db 85402 AATACTACATACACAACTCCCAATGATCAGTGGGATCTCCCTCCACACAATA 85343  
 Qy 6364 GTCTGCTACCTCAGGAATATTTTGTCTGTTGATCCTCAGCTATCGTTGTTGAT 6423  
 Db 85342 GTCTGCTACCTCAGGAATATTTTGTCTGTTGATCCTCAGCTATCGTTGTTGAT 85283  
 Qy 6424 GGTCTTCAGAAATTAATGCTCTCTCATCTTAAATGCCCCATGRCATATACACT 6483  
 Db 85282 GGTCTTCAGAAATTAATGCTCTCTCATCTTAAATGCCCCATGRCATATACACT 85223  
 Qy 6484 GAAACAAGATTATACATTAATGATCTAATGAGCCGCAACAAAGAGTACCATTCTT 6543  
 Db 85222 GAAACAAGATTATACATTAATGATCTAATGAGCCGCAACAAAGAGTACCATTCTT 85163  
 Qy 6544 CCTTTGTATAGAGAGAGAGTGTGATGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 6603  
 Db 85162 CCTTTGTATAGAGAGAGAGTGTGATGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 85103  
 Qy 6604 ACCTCTACATGATCTCTCAATCAATCTAATCAAGATGAGGAGATGAGGAGGATC 6663  
 Db 85102 ACCTCTACATGATCTCTCAATCAATCTAATCAAGATGAGGAGATGAGGAGGATC 85043  
 Qy 6664 GCCGACTCCCTGCTGATCTTCAGATCAATCAATCAATCAATCAATCAATCAATCAAT 6723  
 Db 85042 GCCGACTCCCTGCTGATCTTCAGATCAATCAATCAATCAATCAATCAATCAATCAAT 84983  
 Qy 6724 AATCGAAGAGCTTATGATCTGATCAATCGGATGAGGAGGAGGAGGAGGAGGAGGAG 6783  
 Db 84982 AATCGAAGAGCTTATGATCTGATCAATCGGATGAGGAGGAGGAGGAGGAGGAGGAG 84923  
 Qy 6784 GAAAGAAGCTGATTAATGATCAATCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 6843  
 Db 84922 GAAAGAAGCTGATTAATGATCAATCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 84863  
 Qy 6844 CSAAGATGATTAATCAATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6903  
 Db 84862 CSAAGATGATTAATCAATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84803  
 Qy 6904 AGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6963  
 Db 84802 AGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 84743  
 Qy 6964 CTCTCTCTTGAAGCCCTGATCTTACCTCTGTTAACTTGTCTCTCCAGAAATGAA 7023  
 Db 84742 CTCTCTCTTGAAGCCCTGATCTTAACTCTCTGTTAACTTGTCTCTCCAGAAATGAA 84683  
 Qy 7024 GCTGTAAATCTAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7083  
 Db 84682 GCTGTAAATCTAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84623  
 Qy 7084 CTGAGACGGGCTGTAGCCCAATCTGATGTTAATGATCAATCAAGGAGGAGGAGGAGGAG 7143  
 Db 84622 CTGAGACGGGCTGTAGCCCAATCTGATGTTAATGATCAATCAAGGAGGAGGAGGAGGAG 84563  
 Qy 7144 GAAATCTAAGTGAACAACTCTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7203  
 Db 84562 GAAATCTAAGTGAACAACTCTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84503  
 Qy 7204 SGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7263  
 Db 84502 CGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 84443

Qy 7264 AGGACTAGCTGATTTCTTCTAGGCTGATTAAGAAATCCCTTAAGCTAGGTTGGAGGCTGACC 7323  
 Db 84442 AGGACTAGCTGATTTCTTCTAGGCTGATTAAGAAATCCCTTAAGCTAGGTTGGAGGCTGACC 84383  
 Qy 7324 ACATCCACCTTTTAAACAGGAGGCTTGCATTTAGTTCACACCTGACCAATCAGAGAGCTC 7383  
 Db 84382 ACATCCACCTTTTAAACAGGAGGCTTGCATTTAGTTCACACCTGACCAATCAGAGAGCTC 84323  
 Qy 7384 ACTAAATGCTAATTAAGGCAAGAGAGGATTAAGAAATAGCCATCATTTATGTCMTG 7443  
 Db 84322 ACTAAATGCTAATTAAGGCAAGAGAGGATTAAGAAATAGCCATCATTTATGTCMTG 84263  
 Qy 7444 AGAGCAG 7503  
 Db 84262 AGAGCAG 84203  
 Qy 7504 CAACCCCTTTGGGCTCCCTCTCTGATGAGAGCTGTGTTTTCATGCTATTTCACTCT 7563  
 Db 84202 CAACCCCTTTGGGCTCCCTCTCTGATGAGAGCTGTGTTTTCATGCTATTTCACTCT 84143  
 Qy 7564 ATTAATCTTTCAGCTGCR 7582  
 Db 84142 ATTAATCTTTCAGCTGCA 84124  
 RESULT 10  
 AY101583  
 LOCUS 10222 bp DNA linear PRI 11-FEB-2004  
 DEFINITION Homo sapiens isolate 21 endogenous retrovirus HERV-W, ERVW1 locus, allele B, complete sequence.  
 ACCESSION AY101583  
 VERSION AY101583.1 GI:37544401  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 10222)  
 AUTHORS Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnaud,B., Lucotte,G., Duret,L. and Mandrand,B.  
 TITLE The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)  
 PUBMED 14757826  
 REFERENCE 2 (bases 1 to 10222)  
 AUTHORS Mallet,F., Bouton,O. and Oriol,G.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-biowetex, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 cedex 07, France  
 FEATURES  
 source  
 1..10222  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /isolate="21"  
 /isolation\_source="PBMC of individual 21"  
 /db\_xref="taxon:9606"  
 /map="7q21-q22"  
 /chromosome="7"  
 /sex="female"  
 /cell\_type="PBMC"  
 /note="Isolated by PCR: ERVW1 locus, allele B endogenous\_virus: HERV-W"  
 1..780  
 LTR  
 /note="putative mRNA transcript 1; gag-pol pseudogene"  
 join(248..811,2887..9813)  
 mRNA  
 /product="envelope glycoprotein"  
 /note="putative mRNA transcript 2"  
 join(248..811,9248..9813)  
 mRNA  
 /note="putative mRNA transcript 3; no coding region"  
 7812..9428  
 CDS

```
/note="syncytin"
/codon_start=1
/product="envelope glycoprotein"
/protein_id="AAM68162.1"
/db_xref="GI:37544402"
/translation="MALPPIHIFLFTVLLPSTLTAPPCRCMISSSPYQEFWMORP
GNIDAPSRSLSKPTFTFAHTMPNRCYHSATLCMHANTHYMKMINPSCPGGLV
TVCMTYPTQMGSDGQVODAKREKAEVLSQLRVHGTS PYKGLDLSKHETLRT
HTRLVSLFNTLTGLHEVSAONPNCWICPLNRPVYSIPEPOMDNFSTHETLV
LVGLVSNLEITHTSNLCVFSNTYTNOCIRMTPTPOIVCLPSGIPEVCGTSA
YRCINSSSEMCFLSPVPMPTIYTBODIXYVSKRNRPVILPVIAGUCVGAAG
TGIGITTSYTFYKLSOELNGMERVADS.LVITODLNSLAVALONRRRLDLTAE
RGGCTPLGECCCYVNOGSIYVEKVEIKRDIRORAELELNTGPMWLLQMMWILP
FLGPLAIIILLFLFPCIFNLVNFVSRLEAVLQMEPKQSKTKLYRRBLDRPASP
RSDVNDIKGTPPEIRISAAPLRPNSAGSS"
9487.10222
```

ORIGIN  
LTR

```
Query Match      85.0%; Score 6446.4; DB 9; Length 10222;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 6680; Conservative 195; Mismatches 51; Indels 53; Gaps 27;

QY 606 GATGGGAAACGTTCCCGCAAGACAAAAGCCCCCTTAAGACGTATTTGGABAAATTGGGA 665
DB 2886 GATGGGAAACGTTCCCGCAAGACAAAAGCCCCCTTAAGACGTATTTGGGAAATTGGGA 2945
QY 666 MCAATTTGACCCCTCAGACATTAAGAAAAGAAAGCACTATATTCCTTGGAGTGGCGGCTG 725
DB 2946 CCAATTTGACCCCTCAGACATTAAGAAAAGAAAGCACTATATTCCTTGGAGTGGCGGCTG 3005
QY 726 GCACTCCTGAGGAGATTAATTAACAACAATCTTACAGCTAGACCTCTTTTGTAGAA 785
DB 3006 GCACTCCTGAGGAGATTAATTAACAACAATCTTACAGCTAGACCTCTTTTGTAGAA 3065
QY 786 AAGGCAAAATGAGATGAAGTCCCATAGTACAACTTTCTTTTCAATTAAGACAACTCAC 845
DB 3066 AAGGCAAAATGAGATGAAGTCCCATAGTACAACTTTCTTTTCAATTAAGACAACTCAC 3125
QY 846 AATTATGTAAAAAAGTGAATTTATGCCCTACAGGAGCCCTTACAGACTACCTCCCATC 905
DB 3126 AATTATGTAAAAAAGTGAATTTATGCCCTACAGGAGCCCTTACAGACTACCTCCCATC 3185
QY 906 CCAGCATCCCGCACTCCTTCCCAATTAAGAGACCCCTTCAACCCAAATGCTCCA 965
DB 3186 CCAGCATCCCGCACTCCTTCCCAATTAAGAGACCCCTTCAACCCAAATGCTCCA 3245
QY 966 AAGGAGATTAAGACAAAAGGTTAAACAGTGAACCAAGAGTCCCAATTTCCCAATTATGA 1025
DB 3246 AAGGAGATTAAGACAAAAGGTTAAACAGTGAACCAAGAGTCCCAATTTCCCAATTATGA 3305
QY 1026 CCCCTCCCAAGCAGTGGGAGGAAGAAATTCGGGCCCAAGCAGAGTGAATGCVTTTYY 1085
DB 3306 CCCCT-CCAAAGCAGTGGGAGGAAGAAATTCGGGCCCAAGCAGAGTGAATGCVTTTYY 3364
QY 1086 TCTCCGAGACTTAAGCAATAAAAAAGAGCTTAAGTAATCTCAGATAAATCTGATG 1145
DB 3365 TCTCCGAGACTTAAGCAATAAAAAAGAGCTTAAGTAATCTCAGATAAATCTGATG 3424
QY 1146 CTATATGTATGTTTAAACAAGGTTAAGACAAATCTTTGATCTGACATGAGAGATATTA 1205
DB 3425 CTATATGTATGTTTAAACAAGGTTAAGACAAATCTTTGATCTGACATGAGAGATATTA-A 3483
QY 1206 TGTCACTGTCTAATCAGACACTAACCCCAATAGAGAGTGCACACATACTGAGCCCT 1265
DB 3484 TGTCACTGTCTAATCAGACACTAACCCCAATAGAGAGTGCACACATACTGAGCCCT 3543
QY 1266 GAGGTTTGGCGATCTCTGGTATCTCAGTCAAGGTCAATGATANGATGAACAAGAG 1325
DB 3544 GAGGTTTGGCGATCTCTGGTATCTCAGTCAAGGTCAATG--ATAGAGATGAACAAG-AGG 3600
QY 1326 AAAAGAAATGATTTCCCAACAGGCGACAGCAGTTCCTCACTTASACCTCATTTGGGAGC 1385
DB 3601 AAAAGAAATGATTTCCCAACAGGCGACAGCAGTTCCTCACTTASACCTCATTTGGG--- 3657
```

```
QY 1386 ACAGAAATCAGTAACATGGGAGATTGGTGCAGACATTTGCTAATCTGTGCTCTASA 1445
DB 3658 ACACAGAAATCAGAAACAT-GGAGATTGGTGTGCAGACATTTGCTAATCTGTGCTCTAGA 3716
QY 1446 GAGCTAAGGAAAACTASGAGAAARCTAAGAAATTACTCAATGATGTCCACCATACACA 1505
DB 3717 GAGCTAAGGAAAACTAGGAAG-AAGTCAAGAAATTACTCAATGATGTCCACCATACACA 3775
QY 1506 GGGGAAAGGAAAGAAATCTCACTGCTCTTCTGAGAGACCTAAGGAGCATTGGAAGC 1565
DB 3776 -GGGAAAGGAAAGAAATCTCACTGCTCTTCTGAGAGACCTAAGGAGCATTGGAAGC 3834
QY 1566 GTGCTCTCTGTCACTGACCTCTCTTGAAGGCCCACTAATCTTAAAGCGTAAGTTTCA 1625
DB 3835 GTGCTCTCTGTCACTGACCTCTCTTGAAGGCCCACTAATCTTAAAGCGTAAGTTTCA 3894
QY 1626 CTCAGTCAGCTGAGACATTAG-AAAAAATTCAAAAGTCTGCCGTAGGCCCGAGCAAA 1684
DB 3895 CTCAGTCAGCTGAGACATTAGAAAAAAATCTTCAAAAGTCTGCCGTAGGCCCGAGCAAA 3954
QY 1685 ACTTAAGAACCTTAATTGAACCTTGCAACTTGCGTTTATTAATAGATCAGAGAGC 1744
DB 3955 ACTTAAGAACCTTAATTGAACCTTGCGCAACTCGGTTTTTATTAATAGATCAGAGAGC 4014
QY 1745 AGGCGGAAACAGACCAACGGGATTTAAAAAAAAGGCCACGCTTTAGTCATGACCCTGAG 1804
DB 4015 AGGCGGAAACAGACCAACGGGATTTAAAAAAAAGGCCACGCTTTAGTCATGACCCTGAG 4074
QY 1805 CAACTGGAATTTGAGGCTCTGAAAAAGGAAAAAGTGGCAATTGAATGCCTAATAGG 1864
DB 4075 CAACTGGAATTTGAGGCTCTGAAAAAGGAAAAAGTGGCAATTGAATGCCTAATAGG 4134
QY 1865 GCTTGCTTCAAGTGCAGGCTTCAAGAGCACTTTAAAAAAGATTGTCCAAGTGAAGTAA 1924
DB 4135 GCTTGCTTCAAGTGCAGGCTTCAAGAGCACTTTAAAAAAGATTGTCCAAGTGAAGTAA 4194
QY 1925 CGGCCCCCTTGATCAGGCCCTTAATTCAAGGGAATCACTGAAAGGCCCATGCCCCAGG 1984
DB 4195 CGGCCCCCTTGATCAGGCCCTTAATTCAAGGGAATCACTGAAAGGCCCATGCCCCAGG 4254
QY 1985 GGAACAAGGCTTTTGAAGTCAGAAAGCACTAACAGATGATCCAGACAGACGAGCTAGG 2044
DB 4255 GGAACAAGGCTCTCTGATCAGAAAGCACTAACAGATGATCCAGACAGACGAGCTAGG 4314
QY 2045 TGCCTGGGCAAGCGGCATCCATGCCATCACTTCAAGAGCCCTGGGATATGCTTACC 2104
DB 4315 TGCCTGGGCAAGCGGCATCCATGCCATCACTTCAAGAGCCCTGGGATATGCTTACC 4374
QY 2105 ATTGAAGGCAAGGAGTGTCTCCCTGAGCACTGGTGGGCTCTTAAATCTTAACTCTTC 2164
DB 4375 ATTGAAGGCAAGG-AAGTGTCTCTGAGCACTGGTGGGCTCTTAAATCTTAACTCTTC 4433
QY 2165 TGTCCCGGACCACTGTCTCTCAGATCTGTCACTATTCTGAGGGGCTCCTTAAGCGGCA 2224
DB 4434 TGTCCCGGACCACTGTCTCTCAGATCTGTCACTA-TCAGAGGGGCTC-TAAGCGGCA 4491
QY 2225 GTCACTAGAACTTTTCCACGCACTAAGTTAAGATCGGGAGCTTTATCTTTTAC 2284
DB 4492 GTCACTAGAACT-TCCTCCACGCACTAAGTTATG-ACTGGGAGGCTTTATCTTTTAC 4549
QY 2285 ATGCTTTTCTAATTAATGCTTGAAGGCCCACTAATCTTGTAGGAGAGACATCTTGACA 2344
DB 4550 ATGCTTTTCTAATTAATGCTTGAAGGCCCACTAATCTTGTAGGAGAGACATCTTGACA 4609
QY 2345 AAGCAGGGGCACTTAATACCTGAACATAGAGAAAGAACACCGCTTTGTGTGCCCTG 2404
DB 4610 AAGCAGGGGCACTTAATACCTGAACATAGAGAAAGAACACCGCTTTGTGT-CCCTGG 4668
QY 2405 CTTGAGAAAGAAATTAATCTTAAGTCTGGGCAACAGAAAGCAATATGAGACAGCCAAA 2464
DB 4669 CTTGAGAAAGAAATTAATCTTAAGTCTGGGCAACAGAAAGCAATATGAGACAG-CAAA 4727
```

QY	2465	GAATGCCCGTCCTGTTCAAGTTAACTPAAAGATTCCATCTCTTCCTCCCTACCAAAAGCA	2524
QY	4728	GAATGCCCGTCCTGTTCAAGTTAACTPAAAGATTCCATCTCTTCCTCCCTACCAAAAGCA	4787
QY	2525	GTATCCCCCTCAGACCCAAAGGCCCAACAAGATTCCAAAAGATTGTTAAAGCACTPAAAGC	2584
Db	4788	GTATCCCCCTCAGACCCCAAGGCCCAACAAGACTCCAAAAGATTGTTAAAGCACTPAAAGC	4847
QY	2585	CCAAAGCTTAGTAAACCATGCAATPACTCCCTGCAAGTAATTCCGTAAGTAAGAGAGG	2644
Db	4848	CCAAAGCTTAGTAAACCATGCAAGTAACCTCCGCAAGTAAGTCC-----AATTTAAGAG	4900
QY	2645	CACAGAAACCCAGTGGACATGGAGGGTTAGTCAAGATCTCAGAGATTATCAATGAGAGC	2704
Db	4901	TACAGAAACCCCAACAGACAGTGGAGGTTAGTCAAGATCTCAGAGATTATCAATGAGAGC	4958
QY	2705	CGTTGTCCTTTATACCAGCTGTAACCTGACCTTATATCTGTGATTTCCCAAAATACAGA	2764
Db	4959	TGTTGTCCTTTATAGCCAGCTGTAACCTGACCTTATATCTGTGCTTCCCAAAATACAGA	5018
QY	2765	GGAAGCAGAGTGTTTACAATCCCTGGAACCTTTAGAGATGCTTCTTGTGCAATCCCTGTACA	2824
Db	5019	GGAAGCAGAGTGTTTACAATCCCTGGAACCTTTAGAGATGCTTCTTGTGCAATCCCTGTACA	5078
QY	2825	TCTGACTCTCAATCTTGTGTTTGCCTTTGAAAGATCTTCAAAACCAACATCTCACTCAC	2884
Db	5079	TCTGACTCTCAATCTTGTGTTTGCCTTTGAAAGATCTTCAAAACCAACATCTCACTCAC	5138
QY	2885	CTGGACTRTTTTACCCCAAGGGTTCAAGGATAGTCCCATCTATTGTGGCCAGGCAATTAGC	2944
Db	5139	CTGGACTRTTTTACCCCAAGGGTTCAAGGATAGTCCCATCTATTGTGGCCAGGCAATTAGC	5198
QY	2945	CCAAGACTTGGAGYCAATMTATACCTGGACATCTTGTCTCTGRTGATGATGATTT	3004
Db	5199	CCAAGACTTGGAGYCAATMTATACCTGGACATCTTGTCTCTGRTGATGATGATTT	5256
QY	3005	ACTTTTGGCGCYCTRTTTCAGAAACCTTGTGCATCAACGCCAAGGCTCTTMAATTT	3064
Db	5257	ACTTTTGGCGCGCYCTRTTTCAGAAACCTTGTGTGCATCAACGCCAAGGCTCTTMAATTT	5316
QY	3065	CCCTGCGYACCTGTGGCTTACAWGGTTTCCAAACGAPARCTGACCTGCTCAGACAGAGT	3124
Db	5317	CCCTGCGYACCTGTGGCTTACAWGGTTTCCAAACGAPARCTGACCTGCTCAGACAGAGT	5376
QY	3125	TAAATACTTAGACATPAAATTTATCCAAAGGACACARGGCCCTCAGTGAAGAAATATCCA	3184
Db	5377	-----TACTTAAAGGCTPAAATTTATCCAAAGGACACARGGCCCTCAGTGAAGAAATATCCA	5432
QY	3185	GCCCTATATCGGCTTATCTCTCATCTCAGAAACCTCTAAAGCACTAAGRGARTCTTGGGCT	3244
Db	5433	GCCCTATATCGGCTTATCTCTCATCTCAGAAACCTCTAAAGCACTAAGRGARTCTTGGGCT	5492
QY	3245	AAAYAGYTTCTGCCGAATATGGAATCCCAAGTGTGGRAATATGACGAGYCATTAATTA	3304
Db	5493	AAAYAGYTTCTGCCGAATATGGAATCCCAAGTGTGGRAATATGACGAGYCATTAATTA	5551
QY	3305	CATPATAATTAAGAAACCTCAGAAAGCAATACCATTTARTATAGATGAYAMCTGAGYMR	3364
Db	5552	CATPATAATTAAGAAACCTCAGAAAGCAATACCATTTARTATAGATGAYAMCTGAGYMR	5611
QY	3365	AAAGTGCTTTCCAGGCGCCCTTAAAGAGGCTTTAAACCAAGYCCCAAGTGTAAAGTTGGC	3424
Db	5612	AAAGTGCTTTCCAGGCGCC-----TAAACCAAGGCGCCAGAGTGTAAAGTTGGTCC	5657
QY	3425	AAACGGGCAAGACTTTTSTTYATAATRTCAAGAAAATAAGRAVAACTCTGAGAGTCC	3484
Db	5658	AAACGGGCAAGACTTTTCTTCAATAGTCAACAG-AAAAACAGAAATTAAGTCTTGAAGAGTCC	5716
QY	3485	TTACACAGATCCBAGAGYAGAGCTTGCACACCTGTGGCACTGCTGACCTAAGGAAATTAATG	3544
Db	5717	TTACACAGATCCBAGAGGATGAGCTTGCACACCTGTGGCACTGCTGACCTAAGGAAATTAATG	5776
QY	3545	TATGTGGCAAAAGGTTGTCATCTGTTTATYAGGGTATGAGTGACATGACAGTGTATGAT	3604

Db	5777	TAATGGCAAAAGGTTGACCTCATTTGTTTAACGGTAAGTGTGTGGCAATGACATCTTAAGTAT	5836
Qy	3605	CTGAACCACTTAAAAATATATACAGGRRAGAAGATCTTACTGTGTGGACATCTCATGAKGTGA	3664
Db	5837	CTGAACCACTTAAAAATATATACAGGAAAGAATCTTACTGTGTGGACATCTCATGATGGA	5896
Qy	3665	ATNGCACTTCTCATCTGCTAAAAGGACCTTGTGGCTGTCAACAACGTTTACTTTAAATPTC	3724
Db	5897	ATNGCACTTCTCATCTGCTAAAAGGACCTTGTGGCTGTCAACAACGTTTACTTTAAATPTC	5956
Qy	3725	AGGCTCTATTTACTTGAARGGACAGTGCRCBACTGTGCATCTTGTCAACTCTTAACCCAG	3784
Db	5957	AGGCTCTATTTACTTGAARGGACAGTGCRCBACTGTGCATCTTGTCAACTCTTAACCCAG	6016
Qy	3785	YNCATTTCTTCCAGCAATGGAAGAAAGATPAAATPAACTGTCAACATTAATTTCTC	3844
Db	6017	CCACATTTCTTCCAGCAATGGAAGAAAGATPAAATPAACTGTCAACATTAATTTCTC	6076
Qy	3845	AAACCTTATGCCACTCGAAGGGGACCTTTTGAARGTTCCTTGAATCCGTGACCTTCAC	3904
Db	6077	AAACCTTATGCCACTCGAAGGGGACCTTTTGAARGTTCCTTGAATCCGTGACCTTCAC	6135
Qy	3905	TTTATATACGATGGAAGTTCCCTTTGTAAGAAAAAGCATTCGAAAAGYGGGGTATGCAATG	3964
Db	6136	TTTATATACGATGGAAGTTCCCTTTGTAAGAAAAAGCATTCGAAAAGTGGGGTATGCAATG	6195
Qy	3965	GTCAGTATATGGAATAATTTGAAGATATCCCTTCACTCCAGAACTTATGTCTTAAGCTTR	4024
Db	6196	GTCAGTATATGGAATAATTTGAAGATATCCCTTCACTCCAGAACTTATGTCTTAAGCTTR	6255
Qy	4025	GCGAAACTAATAGCCCTTCAATKGGGCACTGAAGATTAGAGAAAGBAAAAAGGYYAATATTA	4084
Db	6256	GCGAAACTAATAGCCCTTCACTTGGGCACTGAAGATTAGAGAAAGBAAAAAGGCBAAATTA	6315
Qy	4085	TATACAGACTCTAATATATGCTTAACTAGTCNTCCATGCCCATGMRCAATATGSRAGA	4144
Db	6316	TATACAGACTCTAATATATGCTTAACTAGTCNTCCATGCCCATGCCCATATGSAAGA	6375
Qy	4145	AGGGAATTCCTAACTTTCYAGRGAAACCTTATCMAACATCAGGAAGCCATTAGBARATTA	4204
Db	6376	AGGGAATTCCTAACTTTCYAGRGAAACCTTATCMAACATCAGGAAGCCATTAGBARATTA	6435
Qy	4205	TTATGTGCWCTATACGAACCTTARAGGGGGMAGTCTTACATGCVGGGGTCACTANAAA	4264
Db	6436	TTATGTGCWCTATACGAACCTTARAGGGGGMAGTCTTACATGCGGGGGTCACTANAAA	6495
Qy	4265	GGAAGBAAAAGGAATATAGAGGAATGTCACAAGCAKATATTGAACMAAAAAAGACTCA	4324
Db	6496	GGAAGBAAAAGGAATATAGAGGAATGTCACAAGCAKATATTGAAGCCAAAAAGACTCA	6555
Qy	4325	AGGCAAGACCCTCATTTAGAAATGCTTATTAACCTTCCTTATGATATAGGTAAATCCCTTC	4384
Db	6556	AGGCAAGACCCTCATTTAGAAATGCTTATTAACCAACCCCTGATATAGGTAAATCCCTTC	6615
Qy	4385	CGGGAACCAAGCCCAAGTATCTCAGAGAGAGAAACAAGATGGGGAACCTCAGAGG-CAG	4443
Db	6616	CGGGAACCAAGCCCAAGTATCTCAGAGAGAGAAACAAGATGGGGAACCTCAGAGGACAG	6675
Qy	4444	TTTTCTCCCTCGGGAACGTTAGCCACTGAAGAAGGAAATATCTTTTGCCTGCAACTAT	4503
Db	6676	TTTTCTCCCTCGGGAACGTTAGCCACTGAAGAAGGAAATATCTTTTGCCTGCAACTAT	6735
Qy	4504	CCAATGAAATTACTTAAAAACCTTCATCAAACTTTTCACTTAAGCATGATAGCAACCA	4563
Db	6736	CCAATGAAATTACTTAAAAACCTTCATCAAACTTTTCACTTAAGCATGATAGCAACCA	6795
Qy	4564	TCBAATGGCCAAATCATTTATTTTCTGGAACAAGCCCTTTCAAAATCATTAACACABATXT	4623
Db	6796	TCBAATGGCCAAATCATTTATTTTCTGGAACAAGCCCTTTCAAAATCATTAACACABATXT	6855
Qy	4624	CAGGGCTGGAAXGTGGCCARABAATATCCCTCGCTTAATCCGCAAGCTCTTCAGG	4683

```
Db 6856 CAGGCGCTGTGAAGTGTGCGCAGAGAAATATCCCTGCTTATGCGCAAGCTCTTCAGG 6915
Qy 4684 ARAACAAARAA CAGGSCCAATTACCTGARAARACTGGCAACTGATTTTATCCCAAGAGGCC 4743
Db 6916 AGAACAAAGAACAGGCCATTAACCTGAGAGAGACTGGCACTGATTTTATCCCAAGAGGCC 6975
Qy 4744 AAACCTCAGGAGTTTCACTATCTACTAGTCTGGGTAATACTTTTACAGGGTTGGCAGAG 4803
Db 6976 AAACCTCAGGAGTTTCACTATCTACTAGTCTGGGTAATACTTTTACAGGGTTGGCAGAG 7035
Qy 4804 GCCTTCCCTGTAGAGCAGAAAGGCCCAAGAGTAATTAAGAGCACTAGTTTCACTAATA 4863
Db 7036 GCCTTCCCTGTAGAGCAGAAAGGCCCAAGAGTAATTAAGAGCACTAGTTTCACTAATA 7095
Qy 4864 ATTCCAGATTGCGACTTCCCGAGGCTTACAGAGTGAACAATPAGCCCTGCTTTCCAGGCC 4923
Db 7096 ATTCCAGATTGCGACTTCCCGAGGCTTACAGAGTGAACAATPAGCCCTGCTTTCCAGGCC 7155
Qy 4924 ACAGTAAACCCAGGAGTATCCCAAGGCGTTAGGTATACGATATCACTTACACTGCGCTGA 4983
Db 7156 ACAGTAAACCCAGGAGTATCCCAAGGCGTTAGGTATACGATATCACTTACACTGCGCTGA 7215
Qy 4984 AGGCCAAGTCTCTAAGGGAAGGTGAGAAATGAAATGAATGAATACTCAAGAGCACTTAA 5043
Db 7216 AGGCCAAGTCTCTAAGGGAAGGTGAGAAATGAAATGAATGAATACTCAAGAGCACTTAA 5043
Qy 5044 AAGCAAAACCCAGGAAACCCACCTCAACATGAGCTGTGTTGCGCTATAGCCTTAA 5103
Db 7276 AAGCAAAACCCAGGAAACCCACCTCAACATGAGCTGTGTTGCGCTATAGCCTTAA 7335
Qy 5104 ATCTGCAACTTTCCCAAAAAGCAGAGCTTAGCCCATACGAATGCTGTATGAGAGGCC 5163
Db 7336 ATCTGCAACTTTCCCAAAAAGCAGAGCTTAGCCCATACGAATGCTGTATGAGAGGCC 7395
Qy 5164 TTGATAACCAATGACTTGTGCTTGAGCCCAAGAGCAAGCCAACTTAGTTGAGACATCACT 5223
Db 7396 TTGATAACCAATGACTTGTGCTTGAGCCCAAGAGCAAGCCAACTTAGTTGAGACATCACT 7455
Qy 5224 CCTTAGCAATATACAAAGTTCTTAAACATTAACAAGAAACCTATCCCGAAGAGG 5283
Db 7456 CCTTAGCAATATACAAAGTTCTTAAACATTAACAAGAAACCTATCCCGAAGAGG 7515
Qy 5284 GAAAGAACTATTCACCCMWGTGACATGTAATTAGTCAAGTCCCTTCYCTCTAATCCC 5343
Db 7516 GAAAGAACTATTCACCCMWGTGACATGTAATTAGTCAAGTCCCTTCYCTCTAATCCC 7515
Qy 5344 CATCCCTAAGTATCATCTCTGGAGAGAGCCCTACCAAGTCAATTTATTTAATCCCACTGCGG 5403
Db 7576 CATCCCTAAGTATCATCTCTGGAGAGAGCCCTACCAAGTCAATTTAATCCCACTGCGG 7635
Qy 5404 TTAAAGTGTGAGTGAAGTCTTGATATCATCACTTGAGTGAATAATCCCTGATATGCG 5463
Db 7636 TTAAAGTGTGAGTGAAGTCTTGATATCATCACTTGAGTGAATAATCCCTGATATGCG 7695
Qy 5464 CAAAGAACTGAAAAATCAGAGACAAACCTAGACTATTCCTGTGAACCTCTAGAGGATT 5523
Db 7696 CAAAGAACTGAAAAATCAGAGACAAACCTAGACTATTCCTGTGAACCTCTAGAGGATT 7755
Qy 5524 TGGGCTGCTCTTCAAAACAACAACAGAGAGAAAGTAACTAATAATCATAATCCCGATG 5583
Db 7756 TGGGCTGCTCTTCAAAACAACAACAGAGAGAAAGTAACTAATAATCATAAT-CCCGATG 7814
Qy 5584 GSCCTCCCTTATCATATTTTCTCTATAGTSTTTTAACTSTTTTCACTCACTCACTGA 5643
Db 7815 GSCCTCCCTTATCATATTTTCTCTATAGTSTTTTAACTSTTTTCACTCACTCACTGA 7874
Qy 5644 CCCCCTCAATGCGCTGTATGACCAAGTATCCCTTACCAAGAGTTTCTATGAGAGATG 5703
Db 7875 CCCCCTCAATGCGCTGTATGACCAAGTATCCCTTACCAAGAGTTTCTATGAGAGATG 7934
Qy 5704 CAGCGTCCCGGAAATTTGATGCCCCATCCTATAGAGTCTTTTSTAAGGAAACCCCAACC 5763
Db 7935 CAGCGTCCCGGAAATTTGATGCCCCATCCTATAGAGTCTTTTSTAAGGAAACCCCAACC 7994
Qy 5764 TTCACTGCGCAACCAATATGCCCCGCAACTGCTATCACTCTGCGCACTTTTGCATGAT 5823
Db 7995 TTCACTGCGCAACCAATATGCCCCGCAACTGCTATCACTCTGCGCACTTTTGCATGAT 8054
Qy 5824 GCAATATCTCATTTATGAGACAGAAAAATGATTAATCCTAGTTTCTCGAGAGACTTGA 5883
Db 8055 GCAATATCTCATTTATGAGACAGAAAAATGATTAATCCTAGTTTCTCGAGAGACTTGA 8114
Qy 5884 GTCACTGTCTGTGGACTTACTTCAACCAACTGATATGCTGATGCGGAGTGAAGTTCA 5943
Db 8115 GTCACTGTCTGTGGACTTACTTCAACCAACTGATATGCTGATGCGGAGTGAAGTTCA 8174
Qy 5944 GATCAGGCAAGAGAAAAATGTAAGAAGTATCTCCCACTCACCGGGGTACATGCG 6003
Db 8175 GATCAGGCAAGAGAAAAATGTAAGAAGTATCTCCCACTCACCGGGGTACATGCG 8234
Qy 6004 ACCCTAGCCCTTACAAAGAGCTAGATCTCTCAAAACTAGATGAAGACCTCCGTAACCAT 6063
Db 8235 ACCCTAGCCCTTACAAAGAGCTAGATCTCTCAAAACTAGATGAAGACCTCCGTAACCAT 8294
Qy 6064 ACTGCGCTGTGAAGCTATTTAATACCACTCTCACTGCGCTCAATGAGTCTGCGCCAA 6123
Db 8295 ACTGCGCTGTGAAGCTATTTAATACCACTCTCACTGCGCTCAATGAGTCTGCGCCAA 8354
Qy 6124 AACCTACTAATGTTGGAATATGCGTCCCGCTGAACCTTCARGGCATATGTTGCAATGCC 6183
Db 8355 AACCTACTAATGTTGGAATATGCGTCCCGCTGAACCTTCARGGCATATGTTGCAATGCC 8414
Qy 6184 GTACTGGAACAATGGAACAACCTTCAAGCAGAAATTAACACCTTCCGTTTATGATGA 6243
Db 8415 GTACTGGAACAATGGAACAACCTTCAAGCAGAAATTAACACCTTCCGTTTATGATGA 8474
Qy 6244 CTTCTGTTTCCATGATGGAATTAACCATACCTCAACCTCACTGCTGTAAAATTTAGC 6303
Db 8475 CTTCTGTTTCCATGATGGAATTAACCATACCTCAACCTCACTGCTGTAAAATTTAGC 8354
Qy 6304 AATACTAATACGACAAACCAATGATGATGAGGAGGTTAATCCTCCCAACAATA 6363
Db 8535 AATACTAATACGACAAACCAATGATGATGAGGAGGTTAATCCTCCCAACAATA 8594
Qy 6364 GTCTGCTTACCCACAGAAATATTTTGTGTGTGATACCTCAAGCTATCGTTGTTGAT 6423
Db 8595 GTCTGCTTACCCACAGAAATATTTTGTGTGTGATACCTCAAGCTATCGTTGTTGAT 8654
Qy 6424 GGTCTTTCAAGATCTATGCTCTCTCAATTTTATGTCGCCCTATGRCATCTACAT 6483
Db 8655 GGTCTTTCAAGATCTATGCTCTCTCAATTTTATGTCGCCCTATGRCATCTACAT 8714
Qy 6484 GAAACAAGTTTATACGTTATGATATCTAAGCCCGGAAACAAAGAGTAACCATTTCT 6543
Db 8715 GAAACAAGTTTATACGTTATGATATCTAAGCCCGGAAACAAAGAGTAACCATTTCT 8774
Qy 6544 CTTTGTGTTATGAGACAGAGTGTAGGTGCACTAGTACTGGAATTTGGCGGTATCACA 6603
Db 8775 CTTTGTGTTATGAGACAGAGTGTAGGTGCACTAGTACTGGAATTTGGCGGTATCACA 8834
Qy 6604 ACCTTACTACGATTTCTATCAAACTATCTCAAGAACTAATGAGGACATGAAACGGGTC 6663
Db 8835 ACCTTACTACGATTTCTATCAAACTATCTCAAGAACTAATGAGGACATGAAACGGGTC 8894
Qy 6664 GCGCACTCCCTGATCACTTGAAGAATCAACTTATCCCTAGCAACAGTATGCTTCRA 6723
Db 8895 GCGCACTCCCTGATCACTTGAAGAATCAACTTATCCCTAGCAACAGTATGCTTCRA 8954
Qy 6724 AATCGAAGCTTTATGACTTGTCAACCGCTGARAAGAGGAGAACTGTTTATTTTAAAGG 6783
Db 8955 AATCGAAGCTTTATGACTTGTCAACCGCTGARAAGAGGAGAACTGTTTATTTTAAAGG 9014
Qy 6784 GAAAGATGCTTATTTATGTTATCAATTCGGAATGCTGATGAGAAAGTTAAAGAAATT 6843
Db 9015 GAAAGATGCTTATTTATGTTATCAATTCGGAATGCTGATGAGAAAGTTAAAGAAATT 9074
```



QY	6844	CGAGATCGAATTCACAGCTACACAGACAGAGCTTCGAAACACCTGAGACCTCGGGGCTCTCTC	6903
Db	9075	CGAGATCGAATTCACAGCTACACAGAGAGAGCTTCGAAACACCTGAGACCTCGGGGCTCTCTC	9134
QY	6904	AGCGATGGAATGCCCGTGGATTTCTCCCTCTTAAAGACCTCTGAGACCTTAATATATTGCTA	6963
Db	9135	AGCGATGGAATGCCCGTGGATTTCTCCCTCTTAAAGACCTCTGAGACCTTAATATATTGCTA	9194
QY	6964	CTCTCTTTTGGACCCCTGTATCTTTTACCTCTTGTGTAACCTTTGTCTCTCCAGATCGAA	7023
Db	9195	CTCTCTTTTGGACCCCTGTATCTTTTACCTCTTGTGTAACCTTTGTCTCTCCAGATCGAA	9254
QY	7024	GCTGTAAACCTAACATGAGAGCCCAAGATGACGCTGACGACTTAAGATCTTACCGACACC	7083
Db	9255	GCTGTAAACCTAACATGAGAGCCCAAGATGACGCTGACGACTTAAGATCTTACCGACACC	9314
QY	7084	CTGAGACGGGCGGTATAGGCCCAAGATCTGATATGATATGACATTAAGAGGACCCCTCTGAG	7143
Db	9315	CTGAGACGGGCGGTATAGGCCCAAGATCTGATATGATATGACATTAAGAGGACCCCTCTCTAG	9374
QY	7144	GAAATCTCAGCTGACAACTCTTACTAGCCGCCCAATTGACGAGAGAGAGTATGACGCGT	7203
Db	9375	GAAATCTCAGCTGACAACTCTTACTAGCCGCCCAATTGACGAGAGAGAGTATGACGCGT	9434
QY	7204	SGTGGGCAACCTCCCAACAGACACTTATAGTTTCTGTGATGAGTGGGGACCTGAGAGAC	7263
Db	9435	SGTGGGCAACCTCCCAACAGACACTTATAGTTTCTGTGATGAGTGGGGACCTGAGAGAC	9494
QY	7264	AGGATTTAGCTGGATTTTCTTAGGCTGATTAAGATTCCTTTAAGCTTAGTGGGAAAGTATAC	7323
Db	9495	AGGATTTAGCTGGATTTTCTTAGGCTGATTAAGATTCCTTTAAGCTTAGTGGGAAAGTATAC	9554
QY	7324	ACATCCACCTTTAAACAGGGGCTTGCAACTAGTGTACACCTGACCAATGAGAGAGCTC	7383
Db	9555	ACATCCACCTTTAAACAGGGGCTTGCAACTAGTGTACACCTGACCAATGAGAGAGCTC	9614
QY	7384	ACTAAATATGCTATTAAGGCAAGACAGAGGTTAAAGAAATAGCCATCATATTTATGCTG	7443
Db	9615	ACTAAATATGCTATTAAGGCAAGACAGAGGTTAAAGAAATAGCCATCATATTTATGCTG	9674
QY	7444	AGAGACACAGCAGGAGGAGCAATGATCCGGATTTAAACCAAGTTCGAGCGCGCAACGG	7503
Db	9675	AGAGACACAGCAGGAGGAGCAATGATCCGGATTTAAACCAAGTTCGAGCGCGCAACGG	9734
QY	7504	CAACCCCTTTGGGTCCCTCCCTTGTATGAGGAGCTCGTTTTCATGCTATTTCACTCT	7563
Db	9735	CAACCCCTTTGGGTCCCTCCCTTGTATGAGGAGCTCGTTTTCATGCTATTTCACTCT	9794
QY	7564	ATTAAATCTTGCACTGCR 7582	
Db	9795	ATTAAATCTTGCACTGCA 9813	
RESULT 11			
LOCUS	AY101584	10222 bp	DNA linear PRI 11-FEB-2004
DEFINITION	Homo sapiens isolate 79 endogenous retrovirus HERV-W, ERVW1 locus, allele A, complete sequence.		
ACCESSION	AY101584		
VERSION	AY101584.1	GI:37544403	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 10222)		
AUTHORS	Mallet, F., Boulton, O., Prudhomme, S., Cheynet, V., Oriol, G., Bonnaud, B., Lucotte, G., Duret, L., and Mandrand, B.		
TITLE	The endogenous retroviral locus ERVW1 is a bona fide gene involved in hominoid placental physiology		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)		
PubMed	14757826		
REFERENCES	2 (bases 1 to 10222)		

**AUTHORS** Mallet, P., Bouton, O. and Oriol, G.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (06-MAY-2002) Retrovirology Department, UMR 2142  
**FEATURES** CNS-Dioleureux, Ecole Normale Supérieure de Lyon - 46 allée  
d'Italie, Lyon 69364 cedex 07, France  
**source** location/Qualifiers  
1. .10222  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/isolate="79"  
/isolation\_source="BMC of individual 79"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7q21-q22"  
/sex="male"  
/cell\_type="PBMC"  
/note="caucasian  
isolated by PCR; ERVW1 locus, allele A  
endogenous\_virus: HERV-W"  
1. .780  
join(248. .811,2887. .9813)  
/note="putative mRNA transcript 1; gag-pol pseudogene"  
join(248. .811,7585. .9813)  
/product="envelope glycoprotein"  
/note="putative mRNA transcript 2"  
join(248. .811,9248. .9813)  
/note="putative mRNA transcript 3; no coding region"  
7812. .9428  
/note="syncytin"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AAM68163.1"  
/db\_xref="GI:37544404"  
/translation="MALPYHIFLFLVLLPSFTLTAPPCHRCMTSSSPFOFLRMORP  
GNIDAPYSRLSKGTPTFTAHTPNRCYSATLCMHANTHTWTKINSPCGGLGV  
TYVCTVETORGMSDGGVDQARREKHYEIVSOLTQHGNSPYKGLDLSKHELTAT  
HTRVLSIFNTLLTGLHREVSANQNTNCICPLAPRPVYSIPVEQNNNESTELNTSV  
LVGLVSENEMLTNSLTCVFSNTTYTINSQCIRKTPPTQIVCLPSGIFPVGCSA  
YRCINSSBSCMFLSPVLPMTIYTBODLYNVIISKRNRRVILPPIVIGAVGALG  
TGIGGITTSTQFYKLSOELNGDMERYASLDILODLNSLAIVLQNRBALDLTAE  
RGGCLFLBECQYVNOISGIVTEKYEIRDIRORAEELRNTGPGVLSQMPWILP  
FLPDLAIIITLLRGPCLPVLNPFVNSRIEAVLKQNEPMQSKTKIYRPLDRPASP  
RSDVNDIKTTPPEBISANQPLLRPNAGSS"  
9487. .10222

QY 966 AAGAGATAGACAAAGGGTAAACGTAAACAAAGATGCCAATATTTCCCAATTATGA 1025  
 DB 3246 AAGAGATAGACAAAGGGTAAACGTAAACAAAGATGCCAATATTTCCCAATTATGA 3305  
 QY 1026 CCCCTCCCAAGCAGTGGGAGGAAAGAAATTCGGCCAGCCAGATGCTATGCTTTTXX 1085  
 DB 3306 CCCCT-CCAGCAGTGGGAGGAAAGAAATTCGGCCAGCCAGATGCTATGCTTTTXX 3364  
 QY 1086 TCTCCAGACTTAAAGCAATTAAGACAGACTTATGTAATTCAGATTAACCTGATG 1145  
 DB 3365 TCTCCAGACTTAAAGCAATTAAGACAGACTTATGTAATTCAGATTAACCTGATG 3424  
 QY 1146 CTATATTTGTTTATTAAGAGGTTAAGCAATTTGATCTGACATGAGAGATATATA 1205  
 DB 3425 CTATATTTGTTTATTAAGAGGTTAAGCAATTTGATCTGACATGAGAGATATA-A 3483  
 QY 1206 TGTCACTGCTAAATCAAGCACTAACCCCAATGAGAGATGCCAATACTGACGCT 1265  
 DB 3484 TGTCACTGCTAAATCAAGCACTAACCCCAATGAGAGATGCCAATACTGACGCT 3543  
 QY 1266 GAGGTGTTGGCATCTCTGATATCTCACTAGGTCAATGATAGATGACACAGAGG 1325  
 DB 3544 GAGGTGTTGGCATCTCTGATATCTCACTAGGTCAATG-ATAGATGACACAG-AGG 3600  
 QY 1326 AAGANAATGATTTCCCAAGGCAAGCAGCAGTCCAGTCTAACCCTCATTTGGGAGC 1385  
 DB 3601 AAGANAATGATTTCCCAAGGCAAGCAGCAGTCCAGTCTAACCCTCATTTGGG- 3657  
 QY 1386 AAGANAATCAATGATGAGAGATTTGGTGTCTGACAGACTTTGCTAATCTGTGTCTA 1445  
 DB 3658 AAGANAATCAAGACAT-GAGATGTTGTCTGACAGACTTTGCTAATCTGTGTCTA 3716  
 QY 1446 GGAATTAAGGAAAACTAAGAAATCTATGAATTACTCAATGATGCCAATACACA 1505  
 DB 3717 GGAATTAAGGAAAACTAAGAAATCTATGAATTACTCAATGATGCCAATACACA 3775  
 QY 1506 GGGGAAAGGAAAAATCCTACTGCTTTCTGAGAGACTAAGAGAGCAATTGAGAGAC 1565  
 DB 3776 -GGGAAAGGAAAAATCCTACTGCTTTCTGAGAGACTAAGAGAGCAATTGAGAGAC 3834  
 QY 1566 GTGCTCTCTGCACTGACTGCTTTCTGAGAGCCAACTAATCTTAAAGCTTAATATCA 1625  
 DB 3835 GTGCTCTCTGCACTGACTGCTTTCTGAGAGCCAACTAATCTTAAAGCTTAATATCA 3894  
 QY 1626 CTCACTAGCTGACAGACTTAA-AAAAAATCAAAAGTCTGCGTAAAGCCGAGACAA 1684  
 DB 3895 CTCACTAGCTGACAGACTTAAAGAAATCTTCAAAAGTCTGCGTAAAGCCGAGACAA 3954  
 QY 1685 ACTTAAGAAACCTTAATTAAGCTTGGCAATCTGCGTTTATTAATAGATCAAGAGAC 1744  
 DB 3955 ACTTAAGAAACCTTAATTAAGCTTGGCAATCTGCGTTTATTAATAGATCAAGAGAC 4014  
 QY 1745 AGGCGGAAACAGAACAAAGGATTAAGAAAGGCGCACTTATGATCAAGCCCTCAG 1804  
 DB 4015 AGGCGGAAACAGAACAAAGGATTAAGAAAGGCGCACTTATGATCAAGCCCTCAG 4074  
 QY 1805 CAAGTGAATTTGAGAGCTCTGAGAAAGGAAAGGCGCAATTTGAATGCTAATAG 1864  
 DB 4075 CAAGTGAATTTGAGAGCTCTGAGAAAGGAAAGGCGCAATTTGAATGCTAATAG 4134  
 QY 1865 GCTTGTCTCAAGTGGCTTCAAGAGCACTTAAAGAAAGATTTGCTCAAGTGAAGT 1924  
 DB 4135 GCTTGTCTCAAGTGGCTTCAAGAGCACTTAAAGAAAGATTTGCTCAAGTGAAGT 4194  
 QY 1925 CCGGCTCTTGTCTCAAGTGGCTTCAAGAGCACTTAAAGAAAGATTTGCTCAAGTGAAGT 1984  
 DB 4195 CCGGCTCTTGTCTCAAGTGGCTTCAAGAGCACTTAAAGAAAGATTTGCTCAAGTGAAGT 4254  
 QY 1985 GGAAGAAAGTCTTTTGAAGTCAAGAGCACTAAGCAGATGATCAGAGAGAGCTGAAG 2044  
 DB 4255 GGAAGAAAGTCTTTTGAAGTCAAGAGCACTAAGCAGATGATCAGAGAGAGCTGAAG 4314

QY 2045 TGCCTGGGCAAGCGCATCCATGCGCATCAACCTCAAGAGCCCTGGATATGCTGACC 2104  
 DB 4315 TGCCTGGGCAAGCGCATCCATGCGCATCAACCTCAAGAGCCCTGGATATGCTGACC 4374  
 QY 2105 ATTGAGGCGCAAGAGGTTCTCTGAGCACTGCTGCTTTCTTAATCTTACTTTC 2164  
 DB 4375 ATTGAGGCGCAAG-AGTGTCTCTGAGCACTGCTGCTTTCTTAATCTTACTTTC 4433  
 QY 2165 TGTCCGGAACAATGCTCTCCAGATCTGACATTTCTAGGGGGTCCNTAAGCGGCA 2224  
 DB 4434 TGTCCGGAACAATGCTCTCCAGATCTGACATTTCTAGGGGGTCC-ATAAGCGGCA 4491  
 QY 2225 GTCACTAGATCTTTTCCAGCACTAAGTTAATGAAGTGGAGCTTTATCTTTTAC 2284  
 DB 4492 GTCACTAGATCTTTTCCAGCACTAAGTTAATGAAGTGGAGCTTTATCTTTTAC 4549  
 QY 2285 ATGCTTTTCTAATTAATCTTGAAGCCCACTAAGTGTGAAGAGACATTTAGCAA 2344  
 DB 4550 ATGCTTTTCTAATTAATTAATCTTGAAGCCCACTAAGTGTGAAGAGACATTTAGCAA 4609  
 QY 2345 AAGCAGGGGCAATTAATCACTGAACATAGAGAGAAACCCGTTGTGTTGCCCTG 2404  
 DB 4610 AAGCAGGGGCAATTAATCACTGAACATAGAGAGAAACCCGTTGTGT-CCCTG 4668  
 QY 2405 CTGAGAGAGAAATTAATCTGAAGCTGGGCAAGAGAGCAATATGAGACAGCAAA 2464  
 DB 4669 CTGAGAGAGAAATTAATCTGAAGCTGGGCAAGAGAGCAATATGAGACAG-CAAA 4727  
 QY 2465 GAATGCGGCTCTGTTCAAGTTAACTTAAGATTCACCTCTTCCCTCAAGAGCA 3524  
 DB 4728 GAATGCGGCTCTGTTCAAGTTAACTTAAGATTCACCTCTTCCCTCAAGAGCA 4787  
 QY 2525 GTACCCCTCAGACCCAGGCCCCAAGAGATTCAGAAAGATTTGAAGACTTAAGC 2584  
 DB 4788 GTACCCCTCAGACCCAGGCCCCAAGAGATTCAGAAAGATTTGAAGACTTAAGC 4847  
 QY 2585 CCAAGGCTTAAGTAAACATGATTAATCCCTGACATTAATCCGATGATGAGAGAG 2644  
 DB 4848 CCAAGGCTTAAGTAAACATGATTAATCCCTGACATTAATCCGATGATGAGAGAG 4900  
 QY 2645 CACAGAAACCAAGTGAAGTGAAGGTTAATGTCAGAAATCTAAGATTAATGAGAG 2704  
 DB 4901 TACAGAAACCAAGTGAAGTGAAGGTTAATGTCAGAAATCTAAGATTAATGAGAG 4958  
 QY 2705 GGTGCTCTTAATCCAGGCTTACCTAGCCCTTAATCTGAGTTTCCCAATACCA 2764  
 DB 4959 GGTGCTCTTAATCCAGGCTTACCTAGCCCTTAATCTGAGTTTCCCAATACCA 5018  
 QY 2765 GGAAGCAGAGGTTTAACTCTGAGCTTGAAGATCCCTTCTGATCCCTGTACA 2824  
 DB 5019 GGAAGCAGAGGTTTAACTCTGAGCTTGAAGATCCCTTCTGATCCCTGTACA 5078  
 QY 2825 TCTGACCTCAATCTTGTGTTGCTTGAAGATCTTCAAAACCAACATCTCAACTC 2884  
 DB 5079 TCTGACCTCAATCTTGTGTTGCTTGAAGATCTTCAAAACCAACATCTCAACTC 5138  
 QY 2885 CTGACRTTTTAAACCAAGGTTCAAGGATAGTCCCATATTTGCGAGGCAATTAGC 2944  
 DB 5139 CTGACRTTTTAAACCAAGGTTCAAGGATAGTCCCATATTTGCGAGGCAATTAGC 5198  
 QY 2945 CCAAGACTTGAACATTAATCTGAGCACTTGTCTCTGATGAGTGAATTT 3004  
 DB 5199 CCAAGACTTGAACATTAATCTGAGCACTTGTCTCTGATGAGTGAATTT 5256  
 QY 3005 ACTTTTGTGCTCTTCAAGAAACCTTGTGCTCAAGCAACCAAGCTCTTAAATTT 3064  
 DB 5257 ACTTTTGTGCTCTTCAAGAAACCTTGTGCTCAAGCAACCAAGCTCTTAAATTT 5316  
 QY 3065 CCTGCTAACCTGTGCTCAAGGTTTCAAAAGCTGATCTGCTCAAGAGAGT 3124  
 DB 5317 CCTGCTAACCTGTGCTCAAGGTTTCAAAAGCTGATCTGCTCAAGAGAGT 5376  
 QY 3125 TAAATTAATTAAGTAAATTAATCAAAAGCAGAGGCTCTGATGAGAAATTAATCA 3184



QY 5344 CATCCCTGATATACATCCCTGGGAAGACCCCTACCTCATTTTATTTATCCCAACCTGGG 5403  
DB 7576 CATCCCTGATATACATCCCTGGGAAGACCCCTACCTCATTTTATTTATCCCAACCTGGG 7635  
QY 5404 TTAAAGTGTGAGTGTGAGTCTTGGATACATCACTTGAAGTCAAACTCTGATATGTC 5463  
DB 7636 TTAAAGTGTGAGTGTGAGTCTTGGATACATCACTTGAAGTCAAACTCTGATATGTC 7695  
QY 5464 CAAAGAACCTGAAATCCAGAGACAAACCTAGCTATTCCTGGAACCTCTAGAGATT 5523  
DB 7696 CAAAGAACCTGAAATCCAGAGACAAACCTAGCTATTCCTGGAACCTCTAGAGATT 7755  
QY 5524 TGGGCTGCTCTTCAACCAACCAAGAGAAAAGTAACTAAATCAATAATCCCCATG 5583  
DB 7756 TGGGCTGCTCTTCAACCAACCAAGAGAAAAGTAACTAAATCAATAATCCCCATG 7814  
QY 5584 GGCCTCCCTTATATATTTTCTCTATATGTTSTTTTACCTSTTTCACTCTCATGCA 5643  
DB 7815 GGCCTCCCTTATATATTTTCTCTATATGTTSTTTTACCTSTTTCACTCTCATGCA 7874  
QY 5644 CCCCCTCAAGCGCTGATAGACAGAGTCCCTTACCMAGATTCTATGGAAGATG 5703  
DB 7875 CCCCCTCAAGCGCTGATAGACAGAGTCCCTTACCMAGATTCTATGGAAGATG 7934  
QY 5704 CAGCGTCCGGAATATTTGATGCCCATCGTATAGAGTCTTTSTAAGGGAACCCCAAC 5763  
DB 7935 CAGCGTCCGGAATATTTGATGCCCATCGTATAGAGTCTTTSTAAGGGAACCCCAAC 7994  
QY 5764 TTCACTGCCCAACCCATATGCCCCCACTGCTATCATCTGCACTTTTGCATGAT 5823  
DB 7995 TTCACTGCCCAACCCATATGCCCCCACTGCTATCATCTGCACTTTTGCATGAT 8054  
QY 5824 GGAATATCATTTATGAGCAGAGAAAATGATTAATCTAGTTGTCTGAGAGACTTGA 5883  
DB 8055 GGAATATCATTTATGAGCAGAGAAAATGATTAATCTAGTTGTCTGAGAGACTTGA 8114  
QY 5884 GTCACTGTCTGTTGACTTACTTCAACCAACCTGATGCTGATGGGGTGGAGTTCA 5943  
DB 8115 GTCACTGTCTGTTGACTTACTTCAACCAACCTGATGCTGATGGGGTGGAGTTCA 8174  
QY 5944 GATCAGGCAAGAGAAAATGATTAATCTCCCACTCAACCGGGTACATGTC 6003  
DB 8175 GATCAGGCAAGAGAAAATGATTAATCTCCCACTCAACCGGGTACATGTC 8234  
QY 6004 ACCTCTAGCCCTTACAAAGACCTAGATCTCTCAAACTACAGAAACCTCCGTACCAT 6063  
DB 8235 ACCTCTAGCCCTTACAAAGACCTAGATCTCTCAAACTACAGAAACCTCCGTACCAT 8294  
QY 6064 ACTGCTGTGTAAGCTTATTAATCAACCTTCACTGAGCTCATGAGGCTCGAGCCAA 6123  
DB 8295 ACTGCTGTGTAAGCTTATTAATCAACCTTCACTGAGCTCATGAGGCTCGAGCCAA 8354  
QY 6124 AACCTTACTAATCTGTTGATATGCTCCCTGGAACCTTCAAGCCATATGTTCAATCCT 6183  
DB 8355 AACCTTACTAATCTGTTGATATGCTCCCTGGAACCTTCAAGCCATATGTTCAATCCT 8414  
QY 6184 GAACTGAAACAATGGAACAATCTGAGCAAGAAATTAACAACAATTCCTTTTATGAGA 6243  
DB 8415 GAACTGAAACAATGGAACAATCTGAGCAAGAAATTAACAACAATTCCTTTTATGAGA 8474  
QY 6244 CCTCTTGTTCATATGGAATTAACCATACCTCAAACTCAACCTGTGTTAAATTTAGC 6303  
DB 8475 CCTCTTGTTCATATCTGGAATTAACCATACCTCAAACTCAACCTGTGTTAAATTTAGC 8534  
QY 6304 AATATCATATCAACAACAATCTCCATGATCAAGTGGTAACTCTTCCACACAATA 6363  
DB 8535 AATATCATATCAACAACAATCTCCATGATCAAGTGGTAACTCTTCCACACAATA 8594  
QY 6364 GTTCGCTTACCTCAGGAATATTTTGTCTGTGGAATCTCAGCCTATGTTGTTGAAT 6423  
DB 8595 GTTCGCTTACCTCAGGAATATTTTGTCTGTGGAATCTCAGCCTATGTTGTTGAAT 8654

QY 6424 GGCTTTCAGAATCTATGTCCTTCTCATTTTATGTCGCCCTATGRCATCTACAT 6483  
DB 8655 GGCTTTCAGAATCTATGTCCTTCTCATTTTATGTCGCCCTATGRCATCTACAT 8714  
QY 6484 GAACAAGATTTATACAGTTATGATCATATCAAGCCCCGCAACAAAGAGATCACTTCT 6543  
DB 8715 GAACAAGATTTATACAGTTATGATCATATCAAGCCCCGCAACAAAGAGATCACTTCT 8774  
QY 6544 CCTTTGTTATATAGACAGAGTGTAGTGTGACTATGCTATGCTGCAATGGCGTATACA 6603  
DB 8775 CCTTTGTTATATGAGACAGAGTGTAGTGTGACTATGCTGCAATGGCGTATACA 8834  
QY 6604 ACCTTACTCAGTTCTACTACAAACATATCAAGAACTTAAATGGGACATGGAACGGTCT 6663  
DB 8835 ACCTTACTCAGTTCTACTACAAACATATCAAGAACTTAAATGGGACATGGAACGGTCT 8894  
QY 6664 GCCGACTCCGTGTCACTTGCAGAGATCACTTAACTCCCTAGCACAGTATGCTTGA 6723  
DB 8895 GCCGACTCCGTGTCACTTGCAGAGATCACTTAACTCCCTAGCACAGTATGCTTGA 8954  
QY 6724 AATCGAAGCTTTTATGACTTACCTGATACCTGATAGAGGGGAACTTGTATTTTAAAGG 6783  
DB 8955 AATCGAAGCTTTTATGACTTACCTGATACCTGATAGAGGGGAACTTGTATTTTAAAGG 9014  
QY 6784 GAAGATGCTGTTATATGTTAATCAATCGGAATGTCACAGAGAAAGTTAAAGAAAT 6843  
DB 9015 GAAGATGCTGTTATATGTTAATCAATCGGAATGTCACAGAGAAAGTTAAAGAAAT 9074  
QY 6844 CSAAGATCAATCAAGATAGACAGAGAGTTCGAAACATGAGACCTCGGGCTCTCTC 6903  
DB 9075 CGAGATCGAATCAAGATAGACAGAGAGTTCGAAACATGAGACCTCGGGCTCTCTC 9134  
QY 6904 AGCGATGAGTATCCCTGAGTTTCTCCCTTCTTATAGACTCTAGACCTTATATGCTAT 6963  
DB 9135 AGCGATGAGTATCCCTGAGTTTCTCCCTTCTTATAGACTCTAGACCTTATATGCTAT 9194  
QY 6964 CTCCTCTTTGAGACCTGATCTTTTACCTCCCTGTTAATTTTGTCTTCAAGAAATCGAA 7023  
DB 9195 CTCCTCTTTGAGACCTGATCTTTTACCTCCCTGTTAATTTTGTCTTCAAGAAATCGAA 9254  
QY 7024 GCTGTAAATCAACAATGAGAGCCCAAGATGACGTCCAAAGCTTACCGCAGACCC 7083  
DB 9255 GCTGTAAATCAACAATGAGAGCCCAAGATGACGTCCAAAGCTTACCGCAGACCC 9314  
QY 7084 CTGACCGGCTGTATAGCCCAAGATCTATATGATATGACATGAAGGACCCCTCTGAG 7143  
DB 9315 CTGACCGGCTGTATAGCCCAAGATCTATATGATATGACATGAAGGACCCCTCTGAG 9374  
QY 7144 GAATCTCAGCTGACACATCTTACTACGCCCAATTCAGCAGAGAGAGCTTATGAGCGGT 7203  
DB 9375 GAATCTCAGCTGACACATCTTACTACGCCCAATTCAGCAGAGAGAGCTTATGAGCGGT 9434  
QY 7204 SGTGGCCAACCTCCCAACAGACATATGATTTCTGTGAGATGGGGGACTGAGAGAC 7263  
DB 9435 SGTGGCCAACCTCCCAACAGACATATGATTTCTGTGAGATGGGGGACTGAGAGAC 9494  
QY 7264 AGGACTAGCTGAGTTTCTTATGCTGATATGAATCTTAAAGCTTATGAGTGGAGGTTGAC 7323  
DB 9495 AGGACTAGCTGAGTTTCTTATGCTGATATGAATCTTAAAGCTTATGAGTGGAGGTTGAC 9554  
QY 7324 ACATCAGCTTTTAAACAGGGGGCTTGCACCTTATGACACCTGACCATCAGAGGCTC 7383  
DB 9555 ACATCAGCTTTTAAACAGGGGGCTTGCACCTTATGACACCTGACCATCAGAGGCTC 9614  
QY 7384 ACTAAATGCTAATTTAGCAAGAAGAGAGGTTAAAGAAATGACCATATATTTATGCTG 7443  
DB 9615 ACTAAATGCTAATTTAGCAAGAAGAGAGGTTAAAGAAATGACCATATATTTATGCTG 9674  
QY 7444 AGAGCAGCAGAGAGAGAGAGATGATCGGGAATTAATCCCAAGTATTCAGAGCGGCAAGG 7503  
DB 9675 AGAGCAGCAGAGAGAGAGAGATGATCGGGAATTAATCCCAAGTATTCAGAGCGGCAAGG 9734  
QY 7504 CAACCCCTTTGGGTCCTCTTGTATGAGAGCTCTGTTTCAATGCTATTTCACTCT 7563

Db 9735 CAACCCCTTGGTGCCCTTCTGTATGGAGCTCTGTTTCATGATTTCACTCT 9794  
Qy 7564 ATTAATCTTGCACTGCR 7582  
Db 9795 ATTAATCTTGCACTGCA 9813

RESULT 12  
BD221808 10499 bp DNA linear PAT 17-JUL-2003  
LOCUS BD221808  
DEFINITION Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses.  
ACCESSION BD221808.1 GI:33031578  
VERSION BD221808.1  
KEYWORDS JP 2002518051-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 10499)  
AUTHORS Allet, P.M., Perin, J.P. and Rieger, F.  
TITLE Nucleic sequence and deduced protein sequence family with human endogenous retroviral motifs, and their uses  
JOURNAL Patent: JP 2002518051-A/3 25-JUN-2002  
INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE  
COMMENT OS Homo sapiens (human)  
PN JP 2002518051-A/3  
PD 25-JUN-2002  
PF 23-JUN-1999 JP 2000556036  
PR 23-JUN-1998 FR 98/07920  
PI PARRICK M ALLET, JEAN PIERRE PERIN, FRANCOIS RIEGER, PC C12N15/09, A01K67/027, A61K31/711, A61K39/21, A61K48/00, A61P21/00, PC A61P25/00, PC A61P37/06, C07K14/15, C12Q1/68, C12Q1/70, C12N15/00 CC Nucleic sequence and deduced protein sequence family with CC human endogenous CC retroviral motifs, and their uses  
FH Key Location/Qualifiers  
FT source 1..10499  
FT 1..10499  
FT Location/Qualifiers  
FT 1..10499  
FT /organism="Homo sapiens"  
FT /mol\_type="genomic DNA"  
FT /db\_xref="taxon:9606"

FEATURES  
source 1..10499  
Location/Qualifiers  
1..10499  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 84.9%; Score 6436; DB 6; Length 10499;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;  
Qy 606 GATGGGAAAGCTTCCCGCAAGCAAAAAAGCCCTTAAGAGTAATTCGTGAATTTGGGA 665  
Db 2953 GATGGGAAAGCTTCCCGCAAGCAAAAAAGCCCTTAAGAGTAATTCGTGAATTTGGGA 3012  
Qy 666 MCAATTTGACCTCTGACACTTAAGAAAGAAAGCACTTAATTTCTTTCGCACTGCGCGCTG 725  
Db 3013 CCAATTTGACCTCTGACACTTAAGAAAGAAAGCACTTAATTTCTTTCGCACTGCGCGCTG 3072  
Qy 726 GCACTCTGAGGGAGATTAATTAATTAACACATCTTACAGCTGAGCTTTTGTGAAA 785  
Db 3073 GCACTCTGAGGGAGATTAATTAATTAACACATCTTACAGCTGAGCTTTTGTGAAA 3132  
Qy 786 AAGGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAGACAACCTCAC 845  
Db 3133 AAGGCAATGAGTGAAGTGCATTAAGTACAACTTTCTTTTCAATTAGACAACCTCAC 3192  
Qy 846 AATTATGTAAGAAAGTGAATTTATGCTTACAGAAAGCTTCAAGTCTTACTCCCTATC 905  
Db 3193 AATTATGTAAGAAAGTGAATTTATGCTTACAGAAAGCTTCAAGTCTTACTCCCTATC 3252  
Qy 906 CAGCAATCCCGACTCTTCCCGCAATTAATTAAGAGCCCGCTTAACCAAAATGCTCAA 965

Db 3253 CAGCAATCCCGACTCTTCCCGCAATTAATTAAGAGCCCGCTTCAACCAAAATGCTCAA 3312  
Qy 966 AAGGATATGACAAAGGGTAAACAGTGAAACAAAGAGTGCCTAATTTCCCAATTATGA 1025  
Db 3313 AAGGATATGACAAAGGGTAAACAGTGAAACAAAGAGTGCCTAATTTCCCAATTATGA 3372  
Qy 1026 CCCCTCCCAAGCAGTGGAGAGAGAAATTCGGCCCAAGCAGAGTGCATGTCCTTTTYY 1085  
Db 3373 CCCCT-CCAAGCAGTGGAGAGAGAAATTCGGCCCAAGCAGAGTGCATGTCCTTTTTC 3431  
Qy 1086 TCTCCCAAGCTTAAAGCAATTAAGCACTTAGGTAAATTTCTCAGTAATTCCTATG 1145  
Db 3432 TCTCCCAAGCTTAAAGCAATTAAGCACTTAGGTAAATTTCTCAGTAATTCCTATG 3491  
Qy 1146 CTATATGATGTTTAAAGAGGTTGAGCAATTTCTTGAATGAGTGAAGATATTA 1205  
Db 3492 CTATATGATGTTTAAAGAGGTTGAGCAATTTCTTGAATGAGTGAAGATATTA 3550  
Qy 1206 TGTCACTGCTAATAGACACTAACCCCAATGAGAGAGTGCACATTAATGACGCT 1265  
Db 3551 TGTCACTGCTAATAGACACTAACCCCAATGAGAGAGTGCACATTAATGACGCT 3610  
Qy 1266 GAGGTTTGGCAATCTCTGTAATTCAGTCAAGTCAATGAGATGACACAGAG 1325  
Db 3611 GAGGTTTGGCAATCTCTGTAATTCAGTCAAGTCAATGAGATGACAGAG 3667  
Qy 1326 AAGGAAATGATTTCCCAAGGCGAGAGAGTTCAGTCAATGAGAGAG 1385  
Db 3668 AAGGAAATGATTTCCCAAGGCGAGAGAGTTCAGTCAATGAGAGAG 3724  
Qy 1386 ACAGAAATGATTTCCCAAGGCGAGAGAGTTCAGTCAATGAGAGAG 1445  
Db 3725 ACAGAAATGATTTCCCAAGGCGAGAGAGTTCAGTCAATGAGAGAG 3783  
Qy 1446 GAGCTAAGGAAATCTAGAGAAATCTAGAAATCTAGAAATCTAGAAATCTAGAA 1505  
Db 3784 GAGCTAAGGAAATCTAGAGAAATCTAGAAATCTAGAAATCTAGAAATCTAGAA 3842  
Qy 1506 GGGGAAAGGAGAAATCTAGAGAAATCTAGAGAAATCTAGAGAAATCTAGAGAA 1565  
Db 3843 GGGGAAAGGAGAAATCTAGAGAAATCTAGAGAAATCTAGAGAAATCTAGAGAA 3901  
Qy 1566 GTGCTCTCTGTCACCTGACTCTTCTGAAGGCCAATTAATTTAAAGCTTAATTC 1625  
Db 3902 GTGCTCTCTGTCACCTGACTCTTCTGAAGGCCAATTAATTTAAAGCTTAATTC 3961  
Qy 1626 CTCAGTCAAGTGCAGACATTAG-AAAAAATCTTAAGAAAGTGCCTGAGGCCGAGCAA 1684  
Db 3962 CTCAGTCAAGTGCAGACATTAGAAAAAATCTTAAGAAAGTGCCTGAGGCCGAGCAA 4021  
Qy 1685 ACTTAAGAAATCTTAAGAAATCTTAAGAAATCTTAAGAAATCTTAAGAAATCTTA 1744  
Db 4022 ACTTAAGAAATCTTAAGAAATCTTAAGAAATCTTAAGAAATCTTAAGAAATCTTA 4081  
Qy 1745 AGCGGAAAGAGCAAAAGGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 1804  
Db 4082 AGCGGAAAGAGCAAAAGGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTA 4141  
Qy 1805 CAAAGTGAATTTGAGAGCTTGGAAAGGAAAGTGGGCAATTAAGTCCATTAAGG 1864  
Db 4142 CAAAGTGAATTTGAGAGCTTGGAAAGGAAAGTGGGCAATTAAGTCCATTAAGG 4201  
Qy 1865 GCTTGCTTCAAGGCGGCTCAAGAGCACTTTAAAGAAATTAAGTCCATTAAGG 1924  
Db 4202 GCTTGCTTCAAGGCGGCTCAAGAGCACTTTAAAGAAATTAAGTCCATTAAGG 4261  
Qy 1925 CCGCCCTTGTGTCATGCTTATTTCAAGGAAATCACTGAAGGCCCATGCGCCAGG 1984  
Db 4262 CCGCCCTTGTGTCATGCTTATTTCAAGGAAATCACTGAAGGCCCATGCGCCAGG 4321  
Qy 1985 GGAAGAAAGTCTTTGAGTCAAGAGCACTAACAGATGATCCAGAGAGAGATGAGG 2044



Db 4322 GGAAGAGGCTCTGAGTCAGAGCCATTAACAGATGATCCAGCAGCAGACTGAGG 4381  
Qy 2045 TGCGTGGGGCAAGCGCCATCCATGCCATCACCCTCAGAGCGCCGGATATGCTTGACC 2104  
Db 4382 TGCGTGGGGCAAGCGCCATCCATGCCATCACCCTCAGAGCGCCGGATATGCTTGACC 4441  
Qy 2105 ATTGAAGGCCAGGAAGGTTGCTCTGAGACATGTCGGGTCTTCTTAAGTCTTAAGTCTTC 2164  
Db 4442 ATTGAAGGCCAGG-AGGTGTCTCTGAGACATGTCGGGTCTTCTTAAGTCTTAAGTCTTC 4500  
Qy 2165 TGTCCCGGCAACTGTCTCTGAGACATGTCGGGTCTTCTTAAGTCTTAAGTCTTC 2224  
Db 4501 TGTCCCGGCAACTGTCTCTGAGACATGTCGGGTCTTCTTAAGTCTTAAGTCTTC 4558  
Qy 2225 GTCACTAGACTCTTCTTCCAGCCATTAAGTATGAGACTGGAGGCTTAATCTTTTAC 2284  
Db 4559 GTCACTAGACTCTTCTTCCAGCCATTAAGTATG-ACGTGGAGCTTAATCTTTTAC 4616  
Qy 2285 ATGCTTTCTTAATTAATGCTTGAAAGCCCACTACCTTGTAGGAGACATTTCTAGCA 2344  
Db 4617 ATGCTTTCTTAATTAATGCTTGAAAGCCCACTACCTTGTAGGAGACATTTCTAGCA 4676  
Qy 2345 AAGCAGGGGCCATTAACACCTGAGACATGAGAGAGAGAACACCCGTTTGTGTCCTG 2404  
Db 4677 AAGCAGGGGCCATTAACACCTGAGACATGAGAGAGAGAGAACACCCGTTTGTGTCCTG 4735  
Qy 2405 CTGAGAGAGAAATTAATCTCTGAAGCTCTGGGCAACAGAGAGACATTAATGAGAGAG 2464  
Db 4736 CTGAGAGAGAAATTAATCTCTGAAGCTCTGGGCAACAGAGAGACATTAATGAGAGAG 4794  
Qy 2465 GAATCCCGCTCTGTTCAAGTTAACTAAAGATTCACCTCTTCTTCCCTACCAAGAGCA 2524  
Db 4795 GAATCCCGCTCTGTTCAAGTTAACTAAAGATTCACCTCTTCTTCCCTACCAAGAGCA 4854  
Qy 2525 GTACCCCTCTCAAGCCCAAGGCCCAAGAGATTCCTAAAGATTTGAAGACTTAAAGC 2584  
Db 4855 GTACCCCTCTCAAGCCCAAGGCCCAAGAGATTCCTAAAGATTTGAAGACTTAAAGC 4914  
Qy 2585 CCAAGGCTTAATTAATCAATGACATTAATCTGAGACTTAATCTGAGTGGATTTGAGAG 2644  
Db 4915 CCAAGGCTTAATTAATCAATGACATTAATCTGAGACTTAATCTGAGTGGATTTGAGAG 4967  
Qy 2645 CACAGAAACCAAGTGAAGAGGTTAGTGAAGATCTGAGATTAATCAATGAGAGC 2704  
Db 4968 TACAGAAACCAAGTGAAGAGGTTAGTGAAGATCTGAGATTAATCAAT- GAGGC 5025  
Qy 2705 GGTGTCTTTTAATCAAGCTGATCTAGCCCTTAATCTGAGTTCCTCAATTAACAG 2764  
Db 5026 TGTGTCTTTTAATCAAGCTGATCTAGCCCTTAATCTGAGTTCCTCAATTAACAG 5085  
Qy 2765 GGAAGCAGAGTGGTTACASTCTGAGACTTGAAGATGCTTCTTCTGATCCCTGATA 2824  
Db 5086 GGAAGCAGAGTGGTTACASTCTGAGACTTGAAGATGCTTCTTCTGATCCCTGATA 5145  
Qy 2825 TCTGACTCTAATCTTGTGTTGCTTTGAGATTAATCAATCCACATCTCAACTAC 2884  
Db 5146 TCTGACTCTAATCTTGTGTTGCTTTGAGATTAATCAATCCACATCTCAACTAC 5205  
Qy 2885 CTGACTTTTAAACCCCAAGGTTGAGGATTAAGTCCCATTAATTTGGCAGGACTTAC 2944  
Db 5206 CTGACTTTTAAACCCCAAGGTTGAGGATTAAGTCCCATTAATTTGGCAGGACTTAC 5265  
Qy 2945 CCAAGACTTGAAGCTTAATCAATCTGAGACTTGTGCTTCTGATGAGGATTAAT 3004  
Db 5266 CCAAGACTTGAAGCTTAATCTTAATCTGAGACTTGTGCTTCTGATGAGGATTAAT 5323  
Qy 3005 ACTTTGAGCTGCTTAAGAACTTGTGCTTAAGGACCAAGGCTTAAATTT 3064  
Db 5324 ACTTTGAGCTGCTTAAGAACTTGTGCTTAAGGACCAAGGCTTAAATTT 5383  
Qy 3065 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3124  
Db 5384 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5443

Qy 3125 TAAATCTTAGGCTTAATTAATCAAGAGGACCAAGGCTTCAATGAGAAATTAATCA 3184  
Db 5444 ----TACTTAAGGCTTAATTAATCAAGAGGACCAAGGCTTCAATGAGAAATTAATCA 5499  
Qy 3185 GCTTAATCTGCTTAAATCTTCAATCAAGGACCAAGGCTTCAATGAGAAATTAATCA 3244  
Db 5500 GCTTAATCTGCTTAAATCTTCAATCAAGGACCAAGGCTTCAATGAGAAATTAATCA 5559  
Qy 3245 AATAGGCTTCTGCTTAAATCTTCAATCAAGGACCAAGGCTTCAATGAGAAATTAAT 3304  
Db 5560 AATAGGCTTCTGCTTAAATCTTCAATCAAGGACCAAGGCTTCAATGAGAAATTAAT 5618  
Qy 3305 CASTAATTAAGGAACTGAGAAAGCAATCAATTAATTAATGAGAAATTAATCA 3364  
Db 5619 CASTAATTAAGGAACTGAGAAAGCAATCAATTAATTAATGAGAAATTAATCA 5678  
Qy 3365 AATGAGCTTCTGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 3424  
Db 5679 AATGAGCTTCTGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 5724  
Qy 3425 AACAGGCAAGGCTTCTTCTTAAATTAATCAAGGACCAAGGCTTCAATGAGAAATTAAT 3484  
Db 5725 AACAGGCAAGGCTTCTTCTTAAATTAATCAAGGACCAAGGCTTCAATGAGAAATTAAT 5783  
Qy 3485 TTAACAGGCTTCTGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 3544  
Db 5784 TTAACAGGCTTCTGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 5843  
Qy 3545 TAGTGGCAAGGCTTCTGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 3604  
Db 5844 TAGTGGCAAGGCTTCTGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 5903  
Qy 3605 CTGAAGGCTTAAATTAATCAAGGACCAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 3664  
Db 5904 CTGAAGGCTTAAATTAATCAAGGACCAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 5963  
Qy 3665 AYRGATATCTGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 3724  
Db 5964 AYRGATATCTGAGGCTTAAAGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 6023  
Qy 3725 AGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3784  
Db 6024 AGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6083  
Qy 3785 YCNCAATTTCTTCAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 3844  
Db 6084 YCNCAATTTCTTCAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 6143  
Qy 3845 AAACCTATGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 3904  
Db 6144 AAACCTATGAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 6202  
Qy 3905 TTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3964  
Db 6203 TTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6262  
Qy 3965 GTGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4024  
Db 6263 GTGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6322  
Qy 4025 GCAGAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4084  
Db 6323 GCAGAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6382  
Qy 4085 TATAAGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4144  
Db 6383 TATAAGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6442  
Qy 4145 AGGAAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4204  
Db 6443 AGGAAATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 6502

QY	4205	TTATTCGCMGTACAGAAACCTPABAGAGGTGMACTTTACATCCTGCGGGGTGATCANAAA	4264
Dp	6503	TTATTCGCTGACAGAAACCTPAAAGAGGTGCGAGCTTTACATCCTCGGGGTGATCAAAA	6562
QY	4265	GGAAGABRBAAGGAAATASAGRGAAYGCGACAGCAKATATTGGAAGCAAAAAGAGCTGCA	4324
Dp	6563	GGAABGABAAAGGAAATGAAAGAGACCTGCCACAGATATTGAAGCAAAAAGAGCTGCA	6622
QY	4325	AGGAGAGACCTCCATTGAAATGCTTATTAACCTTCCTTAGATAGGGTAACTCCTTC	4384
Dp	6623	AGGAGAGACCTCCATTGAAATGCTTATTAACCAACCCCTAGATAGGGTAACTCCTTC	6682
QY	4385	CGGGAAACCAAGCCCCCACTACTCAGAGAGABAAAACAGAAATGGGGAAACCTCACAGAGG-CAG	4443
Dp	6683	CGGGAAACCAAGCCCCCACTACTCAGAGAGABAAAACAGATAGGGAAACCTCACAGAGAGCG	6742
QY	4444	TTTTCTCCCTCGGGACGGTTAGCCATCGAAGAAAGGAAATACTTTTGCTCGCACTAT	4503
Dp	6743	TTTTCTCCCTCGGGACGGCTAGCCACTGAAAGAAAGGAAATACTTTTGCTCGCACTAT	6802
QY	4504	CCAAATGNAATTACTTTAAACCCCTTCATCAAAACCTTCACTTAGGCAATCGATAGCACCCA	4563
Dp	6803	CCAAATGNAATTACTTTAAACCCCTTCATCAAAACCTTCACTTAGGCAATCGATAGCACCCA	6862
QY	4564	TCABATGSCCAATCATTTATTTACTGGAACCGGCTTTCAAACCTATCAAGCABATATKT	4623
Dp	6863	TCABATGSCCAATCATTTATTTACTGGAACCGGCTTTCAAACCTATCAAGCAGATATAGT	6922
QY	4624	CAGGGCCCTGTAAKTGTGCGCABABAAATATCCCTGCTATGCGCAAGCTCCTTCAGG	4683
Dp	6923	CAGGGCCCTGTAAAGTGTGCGCAGAGAAATATCCCTGCTATGCGCAAGCTCCTTCAGG	6982
QY	4684	ARAAACAABAAACAGGSCCATTAACCTCGABABAACTGSCCACTGATTTTACCACAGAGCC	4743
Dp	6983	AGABCAABAAACAGGSCCATTAACCTCGAGAAAGACTGGCACTGATTTTACCACAGAGCC	7042
QY	4744	AAACCTCAGGGATTCAGTATCTACTAGTCTGGGGTAAATCTTTACGGGGTTGGGCAAG	4803
Dp	7043	AAACCTCAGGGATTCAGTATCTACTAGTCTGGGGTAAATCTTTACCGGTTGGGCGAG	7102
QY	4804	GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGTAAATAAAGGCACTAGTTCATGAATA	4863
Dp	7103	GCCTTCCCTGTAGGACAGAAAGGCCCAAGAGTAAATAAAGGCACTAGTTCATGAATA	7162
QY	4864	ATTCCAGATTTCGACCTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGGCGC	4923
Dp	7163	ATTCCAGATTTCGACCTTCCCGAGGCTTACAGAGTGACAATAGCCCTGCTTTCAGGCGC	7222
QY	4924	ACAGTAAACCCAGAGGATATCCAGGCGTTAGGTATACATATCACTTACATCGGCGCTGA	4983
Dp	7223	ACAGTAAACCCAGAGGATATCCAGGCGTTAGGTATACATATCACTTACATCGGCGCTGA	7282
QY	4984	AGGCCACAGTCTCTCAGGGGAGGTTCAGAAAGAAATGAATACTGAATACTCAATCTAAA	5043
Dp	7283	AGGCCACAGTCTCTCAGGGGAGGTTCAGAAAGAAATGAATACTGAATACTCAATCTAAA	7342
QY	5044	AAGCAAAACCCAGAAACCCACTCACTCATGTGCTGTGTGCTTGTGCTATAGCTTTAAAGA	5103
Dp	7343	AAGCAAAACCCAGAAACCCACTCACTCATGTGCTGTGTGCTTGTGCTTGTGCTTAAAGA	7402
QY	5104	ATCTGCAACTTTCCCCAAABACAGAGACTTAGCCCATAGCAAAATGCTGTATGAAAGGCC	5163
Dp	7403	ATCTGCAACTTTCCCCAAABACAGAGACTTAGCCCATAGCAAAATGCTGTATGAAAGGCC	7462
QY	5164	TTTCATTAACCAATGACCTTGTGCTTGACCCAAACAGGCCAATTAATTTGCAAGCATCACT	5223
Dp	7463	TTTCATTAACCAATGACCTTGTGCTTGACCCAAACAGGCCAATTAATTTGCAAGCATCACT	7522
QY	5224	CCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAAGAACTATCCCTAGAGAAAGG	5283
Dp	7523	CCTTAGCCAAATATCAACAAGTTCTTAAACATTAACAAGAACTATCCCTAGAGAAAGG	7582
QY	5284	GAABAGAACTATTCACCCCMGTGACATGATATGATCAAGTCCCTTCTCTTAATTC	5343

Db	7583	GAAAGAACATATTCACCCCTTGAGACATGGATTAGCAAGTCCCTTCCCTCTAATTC	7642
Qy	5344	CATCCCTAAGATACATCTTGGGAAGAACCTTACCAGTCATTTTATATACCCCAATGCGG	5403
Db	7643	CATCCCTAAGATACATCTTGGGAAGAACCTTACCAGTCATTTTATATACCCCAATGCGG	7702
Qy	5404	TTAAAGGGCGGAGTGGAGCTTGATATATCACTTGAGTCAAAATCCGATATAGTC	5463
Db	7703	TTAAAGGGCGGAGTGGAGCTTGATATATCACTTGAGTCAAAATCCGATATAGTC	7762
Qy	5464	CAAGAACACTGAAAATCCAGAGACAACGCTAGACTATTCCTGTGAACCTCTAGAGATT	5523
Db	7763	CAAGAACAACGAAAATCCAGAGACAACGCTAGACTATTCCTGTGAACCTCTAGAGATT	7822
Qy	5524	TGCGCTGTCTTTCAACACACACAGAGAAAAGTAACTTAAATCAATTAATCCCATG	5583
Db	7823	TGCGCTGTCTTTCAACACACACAGAGAAAAGTAACTTAAATCAATTAATCCCATG	7881
Qy	5584	GSCCGCCCTTATCATATTTTTCTCTKATAGTSTTTTATACCTSTTTCACTCACTGCA	5643
Db	7882	GSCCGCCCTTATCATATTTTTCTCTKATAGTSTTTTATACCTSTTTCACTCACTGCA	7941
Qy	5644	CCCCCTCCATGCGCGCTGTATGACCAAGTAGTCCCTTACCMAGAAGTTCTATGGAAATG	5703
Db	7942	CCCCCTCCATGCGCGCTGTATGACCAAGTAGTCCCTTACCMAGAAGTTCTATGGAAATG	8001
Qy	5704	CAGGCTCCGGAATATTTGATGCCCATCGTATAGAGTCTTSTTAAGGAAACCCCAAC	5763
Db	8002	CAGGCTCCGGAATATTTGATGCCCATCGTATAGAGTCTTSTTAAGGAAACCCCAAC	8061
Qy	5764	TTCACTGCGCCACACCATATGCCCCCGGAACGCTATACACTGCGACACTTTTGATGCAT	5823
Db	8062	TTCACTGCGCCACACCATATGCCCCCGGAACGCTATACACTGCGACACTTTTGATGCAT	8121
Qy	5824	GCAATAACTCATTTATTTGACAGAGAAAATGATTAATCCTAGTTGCTGTGAGACTTGA	5883
Db	8122	GCAATAACTCATTTATTTGACAGAGAAAATGATTAATCCTAGTTGCTGTGAGACTTGA	8181
Qy	5884	GTCACTGTCTGTGGACTTACTTCAACCAACTGATATGTCTGATGGGGGTGGAGTTCAA	5943
Db	8182	GTCACTGTCTGTGGACTTACTTCAACCAACTGATATGTCTGATGGGGGTGGAGTTCAA	8241
Qy	5944	GATCAGGGAAGAAAAACATGTAAAGAAATATATCTCCCACTCACCGSGGTATATGCG	6003
Db	8242	GATCAGGGAAGAAAAACATGTAAAGAAATATATCTCCCACTCACCGSGGTATATGCG	8301
Qy	6004	ACCTTAGACCCCTTCAAGAAGACTAGATCTTCAAACTACATGAAAACCTCCGATCCAT	6063
Db	8302	ACCTTAGACCCCTTCAAGAAGACTAGATCTTCAAACTACATGAAAACCTCCGATCCAT	8361
Qy	6064	ACTGCGCTGTGAAGCTTATTTAATPCCACCTCACTGGGCTTCATGAGTCTCGGCCCAA	6122
Db	8362	ACTGCGCTGTGAAGCTTATTTAATPCCACCTCACTGGGCTTCATGAGTCTCGGCCCAA	8422
Qy	6124	AACCTTACTAATGTGTGATATGCTCCCGCCGMAACTTCAAGCCATATGTTCAATCCCT	6183
Db	8422	AACCTTACTAATGTGTGATATGCTCCCGCCGMAACTTCAAGCCATATGTTCAATCCCT	8481
Qy	6184	GTACCTGAACATGGAACAACCTTGACGACAGAAATPAAACACCACTTCCGTTTATGAGA	6243
Db	8482	GTACCTGAACATGGAACAACCTTGACGACAGAAATPAAACACCACTTCCGTTTATGAGA	8543
Qy	6244	CCTCTGTGTTCCATSTGAAATPAAACCATACCTCAACCTGACCTGTGTAAATTTTATG	6303
Db	8542	CCTCTGTGTTCCATSTGAAATPAAACCATACCTCAACCTGACCTGTGTAAATTTTATG	8601
Qy	6304	AATATCTAATACACAACAACCTCCCAATGCAATGAGTGGGTAAACCTCCCAACAATTA	6363
Db	8602	AATATCTAATACACAACAACCTCCCAATGCAATGAGTGGGTAAACCTCCCAACAATTA	8661
Qy	6364	GTCTGCGTACCTCGAATATTTTGTGCTGTGGTACCTCAGCCTATCGTTGTGAT	6423

Db 8662 GTCGCTACCTCAGGAATATTTTGTCTGTGTACCTCAGCCTATCGTTGTATGAT 8721  
Qy 6424 GGCCTTCAGAAATCATGTGCTTCCTCATTTCTTAGTGCCCGCATGRCATCAACT 6483  
Db 8722 GGGCTTCAGAAATCATGTGCTTCCTCATTTCTTAGTGCCCGCATGRCATCAACT 8781  
Qy 6484 GAACAAATTTATACAGTTATGTATCTAATGAGCCCGCAACAAAGATGCCATCTT 6543  
Db 8782 GAACAAATTTATACAGTTATGTATCTAATGAGCCCGCAACAAAGATGCCATCTT 8841  
Qy 6544 CTTTGTATAGAGAGAGAGTGTAGTCACTAGTACTGGCATTTGGCGTATACA 6603  
Db 8842 CTTTGTATAGAGAGAGAGTGTAGTCACTAGTACTGGCATTTGGCGTATACA 8901  
Qy 6604 ACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGGCATGGAGCGGGTC 6663  
Db 8902 ACCTCTACTAGTTCTACTACAACTATCTCAAGAACTAAATGGGGCATGGAGCGGGTC 8961  
Qy 6664 GCCGACTCCCTGGTCACTTGCAGATCACTTAACCTCCGAGCAGTATGCTTCTA 6723  
Db 8962 GCCGACTCCCTGGTCACTTGCAGATCACTTAACCTCCGAGCAGTATGCTTCTA 9021  
Qy 6724 AATCGAAGCTTTAGACTGTCTAACCCTGAGAGAGGGGAACTGTATTTTATGGG 6783  
Db 9022 AATCGAAGCTTTAGACTGTCTAACCCTGAGAGAGGGGAACTGTATTTTATGGG 9081  
Qy 6784 GAAGAATGCTTATTTATGTTATCAATCCGGAATGCTCACTGAGAAAGTTAAGAAAT 6843  
Db 9082 GAAGAATGCTTATTTATGTTATCAATCCGGAATGCTCACTGAGAAAGTTAAGAAAT 9141  
Qy 6844 CSAGATCGAATACAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6903  
Db 9142 CSAGATCGAATACAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9201  
Qy 6904 ACCCATGAGATGCCCTGAGATTCCTCCCTTTTGAAGCTTACAGCTATATATTTGCTA 6963  
Db 9202 ACCCATGAGATGCCCTGAGATTCCTCCCTTTTGAAGCTTACAGCTATATATTTGCTA 9261  
Qy 6964 CTCCTCTTTGAGACCTGTATCTTTTACCTCCTGTTAACTTTGTCTTCCAGAAATCGAA 7023  
Db 9262 CTCCTCTTTGAGACCTGTATCTTTTACCTCCTGTTAACTTTGTCTTCCAGAAATCGAA 9321  
Qy 7024 GCTGTAAATCTAACAATGAGAGCCCAAGATGAGTCCAGACTAAGATCTAACCAGAGAG 7083  
Db 9322 GCTGTAAATCTAACAATGAGAGCCCAAGATGAGTCCAGACTAAGATCTAACCAGAGAG 9381  
Qy 7084 CTGAGACCGGCTGTATGAGCCAGATGTATGTTAATGACATCAAGAGAGAGAGAGAG 7143  
Db 9382 CTGAGACCGGCTGTATGAGCCAGATGTATGTTAATGACATCAAGAGAGAGAGAGAG 9441  
Qy 7144 GAATCTCAGTGCACAACTCTAAGAGCCCAATTCAGAGAGAGAGAGAGAGAGAGAGAG 7203  
Db 9442 GAATCTCAGTGCACAACTCTAAGAGCCCAATTCAGAGAGAGAGAGAGAGAGAGAGAG 9501  
Qy 7204 SGTGGGCAACCTCCCAACAGCACTTAGATTTTCTGTGAGATGGGGAGCTGAGAGAG 7263  
Db 9502 C-TGGGCAACCTCCCAACAGCACTTAGATTTTCTGTGAGATGGGGAGCTGAGAGAG 9560  
Qy 7264 AGGATCTAGATTTCTTCAAGCTGATTAAGAACTCTTAAGCTTAAAGTGGAAAGTGAAC 7323  
Db 9561 AGGATCTAGATTTCTTCAAGCTGATTAAGAACTCTTAAGCTTAAAGTGGAAAGTGAAC 9620  
Qy 7324 ACATCAGCTTTAAACAGGGGGCTTGAACCTAGTCAACCTGACCAATCAGAGAGAGCT 7383  
Db 9621 ACATCAGCTTTAAACAGGGGGCTTGAACCTAGTCAACCTGACCAATCAGAGAGAGCT 9680  
Qy 7384 ACTAATAATGCTAATTAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7443  
Db 9681 ACTAATAATGCTAATTAAGCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9740  
Qy 7444 AGAGCAGCAG 7503  
Db 9741 AGAGCAGCAG 9800

Qy 7504 CAACCCCTTTGGGATCCCTCCCTTTGTATGAGAGAGCTCTGTTTATGATCTATTTCACTCT 7563  
Db 9801 CAACCCCTTTGGGATCCCTCCCTTTGTATGAGAGAGCTCTGTTTATGATCTATTTCACTCT 9860  
Qy 7564 ATTAATCTTGACATGTCR 7582  
Db 9861 ATTAATCTTGACATGTCR 9879  
RESULT 13  
AX007980 10499 bp DNA linear PAT 06-SEP-2000  
LOCUS AX007980  
DEFINITION Sequence 3 from Patent WO967395.  
ACCESSION AX007980  
VERSION AX007980.1 GI:9995677  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1. Perin, J.P., Rieger, F., and Alliel, P.M.  
Nucleic sequence and deduced protein sequence family with human  
endogenous retroviral motifs and their uses  
Patent: WO 967395-A 29-DEC-1999.  
JOURNAL INST NAT SANTE RECH MED (FR), PERIN, JEAN PIERRE (FR); RIEGER  
FRANCOIS (FR); ALLIEL PATRICK M (FR)  
location/Qualifiers  
FEATURES  
source  
1..10499  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 84.9%; Score 6436; DB 6; Length 10499;  
Best Local Similarity 95.7%; Pred. No. 0;  
Matches 6680; Conservative 195; Mismatches 50; Indels 54; Gaps 28;  
Qy 606 GATGGGAAACGTTCCCGCAAGACAAACCGCCCTAAGACGTATCTGABAAATTGGGA 665  
Db 2953 GATGGGAAACGTTCCCGCAAGACAAACCGCCCTAAGACGTATCTGABAAATTGGGA 3012  
Qy 666 MCAATTTGACCTCAGACATCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 725  
Db 3013 MCAATTTGACCTCAGACATCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3072  
Qy 726 GCACTCCTGAGGAGATTAATTAATTAACACATCTTACAGGTACCTTTGTAGAA 785  
Db 3073 GCACTCCTGAGGAGATTAATTAATTAACACATCTTACAGGTACCTTTGTAGAA 3132  
Qy 786 AAGGCAATGAG 845  
Db 3133 AAGGCAATGAG 3192  
Qy 846 AATTATGTAATAATGTAATTTATGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 905  
Db 3193 AATTATGTAATAATGTAATTTATGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 3252  
Qy 906 CCAGCATCCCGAGCTCTCTCCCAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 965  
Db 3253 CCAGCATCCCGAGCTCTCTCCCAATTAATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 3312  
Qy 966 AAGGAGATGACAAAG 1025  
Db 3313 AAGGAGATGACAAAG 3372  
Qy 1026 CCCCTCCAGCAGTGGGAG 1085  
Db 3373 CCCCT-CCAGCAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3431  
Qy 1086 TCTCCAGACTTAAGCAATTAATAACAGACTTAGTAAATTCAGATTAATCCTGATAG 1145

Db 3432 TCTCCAGACTTAAAGCAATTAATAAAGAGCTTAAGTAATTCAGATTAACCTGATG 3491  
Oy 1146 CTATATGATGTTTTCACAGGGTTAGACAATTCTTGTGATCTGACATGAGAGATATATA 1205  
Db 3492 CTATATGATGTTTTCACAGGGTTAGACAATTCTTGTGATCTGACATGAGAGATATATA-A 3550  
Oy 1206 TGCACTGCTAAATCAAGACCTAACCCCAATGAGAGAGTGCACCATATGCGAGCCT 1265  
Db 3551 TGCACTGCTAAATCAAGACCTAACCCCAATGAGAGAGTGCACCATATGCGAGCCT 3610  
Oy 1266 GAGGTTTGGCGATCTCTGATATCTCAAGAGTCAATGATGATNGATGACACAGAGG 1325  
Db 3611 GAGAGTTTGGCGATCTCTGATATCTCAAGAGTCAATGATGATNGATGACAGAGG 3667  
Oy 1326 AAAAGAAATGATTTCCCAAGGCCAGACAGCAAGTTCACAGTCTAAGCCCTCATTTGGGAC 1385  
Db 3668 AAAAGAAATGATTTCCCAAGGCCAGACAGCAAGTTCACAGTCTAAGCCCTCATTTGGG- 3724  
Oy 1386 ACAAGAAATCAATCAATGAGAGATGCTGACAGACATTTGCTAACTTGTGTGATASAA 1445  
Db 3725 ACAAGAAATCAATCAATGAGAGATGCTGACAGACATTTGCTAACTTGTGTGATASAA 3783  
Oy 1446 GAGCTAAGGAAAATCTASGAAGAAATCTATGAAATTACTCAATGATGTCACCAATACACA 1505  
Db 3784 GAGCTAAGGAAAATCTASGAAGAAATCTATGAAATTACTCAATGATGTCACCAATACACA 3842  
Oy 1506 GGGGAAGGAAAGAAATCTACTGCTTCTTGAGAGACTAAGGAGGCAATTGAGAGG 1565  
Db 3843 -GGGAAGGAAAGAAATCTACTGCTTCTTGAGAGACTAAGGAGGCAATTGAGAGG 3901  
Oy 1566 GTGCTCTCTGTCACTGACTCTTCTGAGGCCAATTAATCTTAAGCGTAAGTTTCA 1625  
Db 3902 GTGCTCTCTGTCACTGACTCTTCTGAGGCCAATTAATCTTAAGCGTAAGTTTCA 3961  
Oy 1626 CTGAGTCACTGAGACTTAG-AAAAAATTCAAAAGTCTGCGTAGGCCGAGGAGAAA 1684  
Db 3962 CTGAGTCACTGAGACTTAGABAAAABCTTCAAAAGTCTGCGTAGGCCGAGGAGAAA 4021  
Oy 1685 ACTTGAAGAACCTTAATGAATTTGGCAACCTTGCTTTTATATATGATCAGAGAGG 1744  
Db 4022 ACTTGAAGAACCTTAATGAATTTGGCAACCTTGCTTTTATATATGATCAGAGAGG 4081  
Oy 1745 AGCGGGAACAGACAAACGGGATTAATAAAGGCCCGCTTAAGTACAGACCTGAGG 1804  
Db 4082 AGCGGGAACAGACAAACGGGATTAATAAAGGCCCGCTTAAGTACAGACCTGAGG 4141  
Oy 1805 CAAGTGAATTTGGAGGCTCTGAAAAGGGAAGTGGGCAAAATTGATGCTTAATAG 1864  
Db 4142 CAAGTGAATTTGGAGGCTCTGAAAAGGGAAGTGGGCAAAATTGATGCTTAATAG 4201  
Oy 1865 GCTTGCTTCAAGTGGGCTTCAAGAGCACTTTAAAAAAGATTGTCAGAGTAAG 1924  
Db 4202 GCTTGCTTCAAGTGGGCTTCAAGAGCACTTTAAAAAAGATTGTCAGAGTAAG 4261  
Oy 1925 CCGGCTTGTGCTCATGCCCCCTTATTTCAAGGGGAATCACTGGAAGGCCCATGCCCCAGG 1984  
Db 4262 CCGGCTTGTGCTCATGCCCCCTTATTTCAAGGGGAATCACTGGAAGGCCCATGCCCCAGG 4321  
Oy 1985 GGAAGAAAGTCTTTTGAAGTGAAGGCACTAACAGATGATCCAGAGCAAGATGAGG 2044  
Db 4322 GGAAGAAAGTCTTTTGAAGTGAAGGCACTAACAGATGATCCAGAGCAAGATGAGG 4381  
Oy 2045 TGCTTGAGGCAAGGCGCATCCATGCAATCACTCAAGAGCCCTGAGTATGCTTGACC 2104  
Db 4382 TGCTTGAGGCAAGGCGCATCCATGCAATCACTCAAGAGCCCTGAGTATGCTTGACC 4441  
Oy 2105 ATTGAAGGCAAGAGGTTGCTCTGAGCACTGCTGAGGCTTCTTAAGTATGCTGCTC 2164  
Db 4442 ATTGAAGGCAAGG-AGGTTGCTCTGAGCACTGCTGAGGCTTCTTAAGTATGCTGCTC 4500  
Oy 2165 TGTCCCGGCAACTGTCTCTCAGATCTGTCACTATTCTGAGGGGCTCMTAAGACGGGCA 2224  
Db 4501 TGTCCCGGCAACTGTCTCTCAGATCTGTCACTA-TCTGAGGGGGTCC-TAAGACGGGCA 4558

Oy 2225 GTCACTAGATCTTTTTCACAGCACTAAGTTATGAACTGAGGAGCTTATCTTTTAC 2284  
Db 4559 GTCACTAGATCTTTTTCACAGCACTAAGTTATGAACTGAGGAGCTTATCTTTTAC 4616  
Oy 2285 ATGCTTTTCAATTAATGCTTGAAGGCCCACTACCTTGTAGGAGACATTTCTAGCAA 2344  
Db 4617 ATGCTTTTCAATTAATGCTTGAAGGCCCACTACCTTGTAGGAGACATTTCTAGCAA 4676  
Oy 2345 AAGCAGGGGCAATTAATCACTGAAATGAGAGAGAAACACCCGTTGTGTCGCCG 2404  
Db 4677 AAGCAGGGGCAATTAATCACTGAAATGAGAGAGAAACACCCGTTGTGTCGCCG 4735  
Oy 2405 CTGAGGAAGGAATTAATCACTGAAATGAGAGAGAAACACCCGTTGTGTCGCCG 2464  
Db 4736 CTGAGGAAGGAATTAATCACTGAAATGAGAGAGAAACACCCGTTGTGTCGCCG 4794  
Oy 2465 GAATGCCGCTCTGTTCAAGTTAACTAAGAAATTCACCTTCTCTTCCCTACCAAGGCA 2524  
Db 4795 GAATGCCGCTCTGTTCAAGTTAACTAAGAAATTCACCTTCTCTTCCCTACCAAGGCA 4854  
Oy 2525 GTAACCCCTCAAGCCCAAGGCCCAACAGATTCCAAAGATTTGTTAAGACTTAAAGC 2584  
Db 4855 GTAACCCCTCAAGCCCAAGGCCCAACAGATTCCAAAGATTTGTTAAGACTTAAAGC 4914  
Oy 2585 CCAAGGCTTAAGTAAACCAATGACATTAATCCCTGACATTAATCCGTAAGGATGAGAGG 2644  
Db 4915 CCAAGGCTTAAGTAAACCAATGACATTAATCCCTGACATTAATCCGTAAGGATGAGAGG 4967  
Oy 2645 CACAGAAACCAAGTGAACAGTGAAGGTTAAGTCAAGATTCAGATTAATCAATGAGG 2704  
Db 4968 TACAGAAACCAAGTGAACAGTGAAGGTTAAGTCAAGATTCAGATTAATCAATGAGG 5025  
Oy 2705 CGTTGCTTTTAAACCAAGTGAACAGTGAAGGTTAAGTCAAGATTCAGATTAATCAATGAGG 2764  
Db 5026 TGTGTTTCTTAATACCAAGTGAACAGTGAAGGTTAAGTCAAGATTCAGATTAATCAATGAGG 5085  
Oy 2765 GGAAGGAGAGTGGTTTAAACCAAGTGAACAGTGAAGGTTAAGTCAAGATTCAGATTAATCAATGAGG 2824  
Db 5086 GGAAGGAGAGTGGTTTAAACCAAGTGAACAGTGAAGGTTAAGTCAAGATTCAGATTAATCAATGAGG 5145  
Oy 2825 TCCTGACTCAATTTCTTGTGCTTGAAGTCAATCAAGTGAAGTCAATCAATCAATCAAT 2884  
Db 5146 TCCTGACTCAATTTCTTGTGCTTGAAGTCAATCAAGTGAAGTCAATCAATCAATCAAT 5205  
Oy 2885 CTGAGCTTTTAAACCAAGTGAACAGTGAAGGTTAAGTCAAGATTCAGATTAATCAATGAGG 2944  
Db 5206 CTGAGCTTTTAAACCAAGTGAACAGTGAAGGTTAAGTCAAGATTCAGATTAATCAATGAGG 5265  
Oy 2945 CCAAGGCTTAAGTAAACCAATGACATTAATCCCTGACATTAATCCGTAAGGATGAGAGG 3004  
Db 5266 CCAAGGCTTAAGTAAACCAATGACATTAATCCCTGACATTAATCCGTAAGGATGAGAGG 5323  
Oy 3005 ACTTTTGTGCTTCAAGTGAAGGCTTGTGCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3064  
Db 5324 ACTTTTGTGCTTCAAGTGAAGGCTTGTGCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5383  
Oy 3065 CTTGCTTCAATGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 3124  
Db 5384 CTTGCTTCAATGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 5443  
Oy 3125 TAAATCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3184  
Db 5444 -TAAATCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 5499  
Oy 3185 GCTTAATGCTTCAATGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 3244  
Db 5500 GCTTAATGCTTCAATGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 5559  
Oy 3245 AATAGGTTTGTGCGGAAATGAGTCCCAAGTATGCGGAAATGAGTCCCAAGTATGCGGAAATGAGT 3304  
Db 5560 AATAGGTTTGTGCGGAAATGAGTCCCAAGTATGCGGAAATGAGTCCCAAGTATGCGGAAATGAGT 5618





Db 7763 CAAAGAACTGAAATCCAGAGACAACGCTAGCTATTCCTGTGAACTCTAGAGATT 7822  
Qy 5524 TGGCGCTGCTCTTCAAAACAACACAGAGGAAAGTAACTAAATCATTAATCCCCATG 5583  
Db 7823 TGGCGCTGCTCTTCAAAACAACACAGAGGAAAGTAACTAAATCATTAAT - CCCATG 7881  
Qy 5584 GSGCTCCCTTATCATATTTTCTCTKTAATGTTSTTTTAAACCTSTTTCACCTCAGTGA 5643  
Db 7882 GCGCTCCCTTATCATATTTTCTCTKTAATGTTSTTTTAAACCTSTTTCACCTCAGTGA 7941  
Qy 5644 CCCCCCTCATGCGCTGTATGACCAAGTAGCTCCCTTACAGAGATTCTATGAGATG 5703  
Db 7942 CCCCCCTCATGCGCTGTATGACCAAGTAGCTCCCTTACAGAGATTCTATGAGATG 8001  
Qy 5704 CAGCGTCCCGGAAATATTGATGCCCATGTATAGAGTCTTTSATAGGAAACCCCAAC 5763  
Db 8002 CAGCGTCCCGGAAATATTGATGCCCATGTATAGAGTCTTTSATAGGAAACCCCAAC 8061  
Qy 5764 TTCACTGCCCCACACCATATAGCCCCGCACTGCTATCACTGCACTCTTGTGATGAT 5823  
Db 8062 TTCACTGCCCCACACCATATAGCCCCGCACTGCTATCACTGCACTCTTGTGATGAT 8121  
Qy 5824 GCAATATCTCATTTTGGACAGAGAAATATGATTAATCTAGTTGTCTGAGAGACTTGA 5883  
Db 8122 GCAATATCTCATTTTGGACAGAGAAATGATTAATCTAGTTGTCTGAGAGACTTGA 8181  
Qy 5884 GTCACTGTGTGGACTTACTTCAACCAACCTGTATGTGTATGAGGGGTGAGTTCA 5943  
Db 8182 GTCACTGTGTGGACTTACTTCAACCAACCTGTATGTGTATGAGGGGTGAGTTCA 8241  
Qy 5944 GATCAGGCGAGAGAAATCATGTAAAGTAATCTCCCACTCACGSGGTATCATGCG 6003  
Db 8242 GATCAGGCGAGAGAAATCATGTAAAGTAATCTCCCACTCACGSGGTATCATGCG 8301  
Qy 6004 ACCTCTAGCCCTCTCAAAAGACAGATCTCTCAAACTATGAAACCTCCGTACCCAT 6063  
Db 8302 ACCTCTAGCCCTCTCAAAAGACAGATCTCTCAAACTATGAAACCTCCGTACCCAT 8361  
Qy 6064 ACTGCGCTGTGAAGCTTATTAATACCACTCATGCGGTCTCATGAGGTCTCGGCCCA 6123  
Db 8362 ACTGCGCTGTGAAGCTTATTAATACCACTCATGCGGTCTCATGAGGTCTCGGCCCA 8421  
Qy 6124 AACCTCTAATCTGTGTATAGCTCCCTCCGAACTTCAAGCATATGTTCAATCCCT 6183  
Db 8422 AACCTCTAATCTGTGTATAGCTCCCTCCGAACTTCAAGCATATGTTCAATCCCT 8481  
Qy 6184 GTACCTGAAACATGGAACAACCTTCAGACAAGAAATMAACCACTTCGGTTTATGAGA 6243  
Db 8482 GTACCTGAAACATGGAACAACCTTCAGACAAGAAATMAACCACTTCGGTTTATGAGA 8541  
Qy 6244 CTTCTTGTTCATATSTGAAATTAACCAATACCTCAACCTCACTGTGTAAATTTAGC 6303  
Db 8542 CTTCTTGTTCATATSTGAAATTAACCAATACCTCAACCTCACTGTGTAAATTTAGC 8601  
Qy 6304 AATATCTAATACCAACCACTCCCAATGCACTCAGTGGGTAACTCTCCGACACAAATA 6363  
Db 8602 AATATCTAATACCAACCACTCCCAATGCACTCAGTGGGTAACTCTCCGACACAAATA 8661  
Qy 6364 GTGCGCTACCTCAGAAATATTTTGTCTGTGTAACCTCAGCTATGTTTGAAT 6423  
Db 8662 GTGCGCTACCTCAGAAATATTTTGTCTGTGTAACCTCAGCTATGTTTGAAT 8721  
Qy 6424 GGCCTTTCAGAAATCTATGTGCTTCTCATTTTATAGTCCCCATAGTCACTTCACT 6483  
Db 8722 GGCCTTTCAGAAATCTATGTGCTTCTCATTTTATAGTCCCCATAGTCACTTCACT 8781  
Qy 6484 GAAACAAGTTTATACAGTTATGTCATATAGAGCCCGCAACAAAGATACCACTTCT 6543  
Db 8782 GAAACAAGTTTATACAGTTATGTCATATAGAGCCCGCAACAAAGATACCACTTCT 8841  
Qy 6544 CTTTGTGTTATAGAGAGAGATGCTAGTGACTAGTACTGCACTTGGCGGTATACCA 6603  
Db 8842 CTTTGTGTTATAGAGAGAGATGCTAGTGACTAGTACTGCACTTGGCGGTATACCA 8901

Qy 6604 ACCTTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGAAACGGGCT 6663  
Db 8902 ACCTTACTCAGTTTCTACTACAAACTATCTCAAGAACTAAATGGGACATGAAACGGGCT 8961  
Qy 6664 GCGGATCCCTGTGTCACTTGCAGATCACTTAATCTCCCTAGACAGATAGTCTTCTGA 6723  
Db 8962 GCGGATCCCTGTGTCACTTGCAGATCACTTAATCTCCCTAGACAGATAGTCTTCTGA 9021  
Qy 6724 AATCGAAGACTTATAGACTTGTCAACCGCTGAAAGAGGGGGAACCTGTTATTTTGGG 6783  
Db 9022 AATCGAAGACTTATAGACTTGTCAACCGCTGAAAGAGGGGGAACCTGTTATTTTGGG 9081  
Qy 6784 GAAAGATGCTTATTAATATGTTAATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAAT 6843  
Db 9082 GAAAGATGCTTATTAATATGTTAATCAATCCGAATCGTCACTGAGAAAGTTTAAAGAAAT 9141  
Qy 6844 CSAGATCGAATACAACTGATAGAGAGAGAGACTTGGAAACCTGGAACCTGGGGCTCTC 6903  
Db 9142 CSAGATCGAATACAACTGATAGAGAGAGAGACTTGGAAACCTGGGGCTCTC 9201  
Qy 6904 AGCGATATGATGCGCTGTGATTTCTCCCTTTTGGAGCTCTGAGCGTATTAATTTGCTA 6963  
Db 9202 AGCGATATGATGCGCTGTGATTTCTCCCTTTTGGAGCTCTGAGCGTATTAATTTGCTA 9261  
Qy 6964 CTCCTCTTGGAGCCCTGTATCTTTTACCTCCTGTTTAACTTGTCTCTTCAGAAATCGAA 7023  
Db 9262 CTCCTCTTGGAGCCCTGTATCTTTTACCTCCTGTTTAACTTGTCTCTTCAGAAATCGAA 9321  
Qy 7024 GCTGTAAACTTACAAATGAGAGCCCAAGATGATGATCAAGACTTAAGTCTACCGAGACC 7083  
Db 9322 GCTGTAAACTTACAAATGAGAGCCCAAGATGATGATCAAGACTTAAGTCTACCGAGACC 9381  
Qy 7084 CTGAGCGGGCTGTATACCCCAAGATGATGATTAAGATCAATCAAGGACCCCTCTGAG 7143  
Db 9382 CTGAGCGGGCTGTATACCCCAAGATGATGATTAAGATCAATCAAGGACCCCTCTGAG 9441  
Qy 7144 GAAATCTCAGCTGACAACTCTTACTACCGCCCAATTCAGCAGAGAGCACTTAAGCGGT 7203  
Db 9442 GAAATCTCAGCTGACAACTCTTACTACCGCCCAATTCAGCAGAGAGCACTTAAGCGGT 9501  
Qy 7204 SGTGCGCAACCTCCCAAGACACTTAAGTTTCTGTTGAGATGGGGGCTAGAGAC 7263  
Db 9502 SGTGCGCAACCTCCCAAGACACTTAAGTTTCTGTTGAGATGGGGGCTAGAGAC 9560  
Qy 7264 AGACCTAGCTGATTTCTTGGGCGATTAAGATCCCTTAAGCTAGTGGGAAAGGTGACC 7323  
Db 9561 AGACCTAGCTGATTTCTTGGGCGATTAAGATCCCTTAAGCTAGTGGGAAAGGTGACC 9620  
Qy 7324 ACATTCACCTTTAAACACGGGGCTTGGCACTTAAGTCAACCTGACCAATCAGAGAGCTC 7383  
Db 9621 ACATTCACCTTTAAACACGGGGCTTGGCACTTAAGTCAACCTGACCAATCAGAGAGCTC 9680  
Qy 7384 ACTAAATGCTTAATTAAGCAAAAGACAGAGGTAAAGAAATGCCAATCATATTAATGCMTG 7443  
Db 9681 ACTAAATGCTTAATTAAGCAAAAGACAGAGGTAAAGAAATGCCAATCATATTAATGCMTG 9740  
Qy 7444 AGACGACGAGAGAGGAGCAATGATCGGGATATTAACCAAGTATTCAGCGCGGACGG 7503  
Db 9741 AGACGACGAGAGAGGAGCAATGATCGGGATATTAACCAAGTATTCAGCGCGGACGG 9800  
Qy 7504 CAACCCCTTTGGGTCCCTCTTGTATGGGAGCTGTGTTTCACTATTTCACTCT 7563  
Db 9801 CAACCCCTTTGGGTCCCTCTTGTATGGGAGCTGTGTTTCACTATTTCACTCT 9860  
Qy 7564 ATTAATCTTGCATCTGCR 7582  
Db 9861 ATTAATCTTGCATCTGCA 9879

RESULT 14  
AY101586  
LOCUS

AY101586 10229 bp DNA linear PRI 11-FEB-2004

DEFINITION Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERVW1  
locus, allele A, complete sequence.  
ACCESSION AY101586  
VERSION AY101586.1 GI:37544407  
KEYWORDS  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM  
REFERENCE  
AUTHORS Pan troglodytes  
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
1 (bases 1 to 10229)  
1 (bases 1 to 10229)  
Mallet, F., Boucon, O., Prudhomme, S., Cheynet, V., Oriol, G.,  
Bonnaud, B., Lucotte, G., Duret, L. and Mandrand, B.  
The endogenous retroviral locus ERVW1 is a bona fide gene involved  
in hominoid placental physiology  
Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)  
14757826  
JOURNAL 2 (bases 1 to 10229)  
PUBMED Mallet, F., Boucon, O. and Oriol, G.  
REFERENCE Direct Submission  
AUTHORS Submitted (06-MAY-2002) Retrovirology Department, UMR 2142  
CNRS-Biometrieux, Ecole Normale Supérieure de Lyon - 46 allée  
JOURNAL d'Italie, Lyon 69364 cedex 07, France  
FEATURES  
source  
1.10229  
Location/Qualifiers  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/isolate="1"  
/isolation\_source="PBMC of chimpanzee 1"  
/db\_xref="taxon:9598"  
/cell\_type="PBMC"  
/note="Isolated by PCR, ERVW1 locus, allele A  
endogenous virus: ERV-W"  
1.779  
join(248..810,2889..9820)  
/note="putative mRNA transcript 1; gag-pol pseudogene"  
join(248..810,7592..9820)  
/product="envelope glycoprotein"  
/note="putative mRNA transcript 2"  
join(248..810,9255..9820)  
/note="putative mRNA transcript 3; no coding region"  
7819..9435  
/note="bryncycin"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AA068165.1"  
/db\_xref="GI:37544408"  
/translation="MALPYHIFLFTVILSFLLTAPPPRCMTSSPFVCEFLRMORP  
GNIDAPSTRFSKGFPTFAHTHMPNCHSATLCHNANTHTWTKIIMPSCGGLGV  
TVCWTYFTHTGSDGGVODARERKIVEIQLTRVHSRSPYKGLDLSKLETLRT  
HTRVLSFHTTLTGLHEVSAQNPICMLPLNRPYVSIIPVEQNNSTENTTSV  
LVGLPVSNEITHSTNLTCVFSNTDTNSOCIRHVPPTQIVCLPSGIFVCGTSA  
YRLNGSSSMCFSLFIVBPMTIYBODLYNVVSKRPKRPILPPIVGAVALGALG  
TGIGITITSTOORYKLSDELNDMERVADSLVTLDDOLNSLAVALONRAIDLITAE  
RGRTLFAGECCIVNOSGIVTERKIEKRIORAEELRNTGFWGLSOWMPILP  
FUGPLAAILLLFGPCIFNLLVNFVKSRIEAVKIQMBRKMSKTKIYRRPDRPAS  
RSDVNDIGTPEEELITAOPLRPNSAGSS"  
9494..10229  
LTR  
ORIGIN  
Query Match 83.7%; Score 6346.4; DB 9; Length 10229;  
Best Local Similarity 94.9%; Pred. No. 0;  
Matches 6625; Conservative 194; Mismatches 110; Indels 52; Gaps 28;  
QY 606 GATGGGAACGTTCCCGGACAGACAAAGCCCGTAAGAGCTATTCTGGAAATTTGGGA 665  
DB 2888 GATGGGAACGTTCCCGGACAGACAAAGCCCGTAAGAGCTATTCTGGAAATTTGGGA 2947  
QY 666 MCAATTGACCTCAGACACTAAGAAAGAAAGACTTATATCTTCTGAGTGGCCGCTG 725  
DB 2948 CCAATTTGACCTCAGACACTAAGAAAGAAAGACTTATATCTTCTGAGTGGCCGCTG 3007  
QY 726 GCATCTCTGAGGAAGTATAATTATAACACCATCTTACAGTGAAGCTTTTGTAGAA 785

DB 3008 GCATCTCTGAGGAAGTATAATTATAACACCATCTTACAGTGAAGCTTTTGTAGAA 3067  
QY 786 AAGGCAATGAGATGAGATGCCATTAAGTACAACTTTCTTTTCAATTAAGACAACTCAC 845  
DB 3068 AAGGCAATGAGATGAGATGCCATTAAGTACAACTTTCTTTTCAATTAAGACAACTCAC 3127  
QY 846 AATTATGTAATAAGTGTGATTTATGCGCTACAGAAAGCTTCAAGAGTCTACCTCCATC 905  
DB 3128 AATTATGTAATAAGTGTGATTTATGCGCTACAGAAAGCTTCAAGAGTCTACCTCCATC 3187  
QY 906 CCAGCATCCCC--GACTCTTCCCACTTATATAGAACCCCTTCAACCAATGCTCC 963  
DB 3188 CCAGCATCCCCCTGACTCTTCCCACTTATATAGAACCCCTTCAACCAATGCTCC 3247  
QY 964 AAAAGGATGATGACAAAGGATTAACAGGAACCAAGATGCAATATTTCCCAATAT 1023  
DB 3248 AAAAGGATGATGACAAAGGATTAACAGGAACCAAGATGCAATATTTCCCAATAT 3307  
QY 1024 GACCCCTCCCAAGCAGTGGAGAGAGAAATTCGCGCCAGCCAGAGTGCATGCTTTT 1083  
DB 3308 GACCCCT-CCAGCAGTGGAGAGAGAAATTTGCGCCAGCCAGAGTGCATGCTTTT 3366  
QY 1084 YTTCTCCAGACTTAAAGCAATTAAGCAAGACTTATAGTAAATTCAGATAATCTGAT 1143  
DB 3367 TCTCTCTCAGACTTAAAGCAATTAAGCAAGACTTATAGTAAATTCAGATAATCTGAT 3426  
QY 1144 GGCATATATGATGTTTAAAGGTTAGACAAATCTTGTATCTGACATGAGAGATTA 1203  
DB 3427 GGCATATATGATGTTTAAAGGTTAGACAAATCTTGTATCTGACATGAGAGATTA 3486  
QY 1204 TATGTCACTGTAAATCAGACACTAACCCCAATGAGAGAGTGCACCATTAATCTGAC 1263  
DB 3487 -ATGTCACTGTAAATCAGACACTAACCCCAATGAGAGAGTGCACCATTAATCTGAC 3545  
QY 1264 CTGAGGTTTGGCATCTCTGTATCTCAGTCAAGTCAATGATANGATGACACAGAA 1323  
DB 3546 CCAGAGTTTGGCATCTCTGTATCTCAGTCAAGTCAATGATANGATGACACAGAA 3602  
QY 1324 GGAAGGANAATATCCCAAGGCGCAGACAGACATGCTCACTCACTCACTTGGG 1383  
DB 3603 GGAAGGANAATATCCCAAGGCGCAGACAGACATGCTCACTCACTCACTTGGG- 3661  
QY 1384 ACACAGAAATCAGTAACATGGAAGATGTTGCTGAGACATTTGCTCACTTGTGCTAS 1443  
DB 3662 --ACACAGAAATCAGTAACATGGAAGATGTTGCTGAGACATTTGCTCACTTGTGCTAG 3718  
QY 1444 AAGACTAAGGAAAATCTASGAAAGAAATCTAAGATTACTCAATGATGCCACATACA 1503  
DB 3719 AAGACTAAGGAAAATCTAAGAAAG--AAGCTAATGAATTACTCAATGATGCCACATACA 3777  
QY 1504 CAGGGGAAAGGAAAATCTCTAATGCTCTTCTGAGAGACTAAGGAGGCACTTGGAA 1563  
DB 3778 CA-GGAAAGGAAAATCTCTAATGCTCTTCTGAGAGACTAAGGAGGCACTTGGAA 3836  
QY 1564 GCGTCCCTCTGCTGACCTGACTCTTCTGAGAGGCAATATCTTAAGCGTAAGTTAT 1623  
DB 3837 GCGTCCCTCTGCTGACCTGACTCTTCTGAGAGGCAATATCTTAAGCGTAAGTTAT 3896  
QY 1624 CACTCAGTCACTGACATTAAG-AAAAAATCTCAAAAGTCTGCGTAGGCGCGAGCA 1682  
DB 3897 CACTCAGTCACTGACATTAAGAAAAAATCTCAAAAGTCTGCGTAGGCGCGAGCA 3956  
QY 1683 AAACCTTGAAGAACCTTATGAAGCTTGGCAACTGCTTTTATATAAGATCAGAGAGA 1742  
DB 3957 AAACCTTGAAGAACCTTATGAAGCTTGGCAACTGCTTTTATATAAGATCAGAGAGA 4016  
QY 1743 GCAGCGGAAACAGACAAACGGGATTAAGAAAAAGGCGCAGCTTATGATGACCTCA 1802  
DB 4017 GCAGCGGAAACAGACAAACGGGATTAAGAAAAAGGCGCAGCTTATGATGACCTCA 4076  
QY 1803 GGCAGGTGAGCTTTGAGGCTCTGGAAGGAGAAAGCTGGGCAATTTGAATGCTTAATA 1862  
DB 4077 GGCAGGTGAGCTTTGAGGCTCTGGAAGGAGAAAGCTGGGCAATTTGAATGCTTAATA 4136

OY 1863 GGGCTTGCTTCAGATGCGGTCTTACAAAGACATTTTAAAAAGTTGTCCAGTAGAAGTA 1922  
DB 4137 GGGCTTGCTTCAGATGCACTTACAAAGACATTTGAAAAAGTTGTCCAGTAGAAGTA 4196  
OY 1923 AGCGCGCCCTTGCTTCAGATGCGGTCTTCAAGGGAATCATCTGAAAGGCCCATCTGCGCCA 1982  
DB 4197 AGCGCGCCCTTGCTTCAGATGCGGTCTTCAAGGGAATCATCTGAAAGGCCCATCTGCGCCA 4256  
OY 1983 GGGGCAAAAGTCTTTTGAATCAAGAACCACTAACAGATGATCCAGCAGCAGGACTGAG 2042  
DB 4257 GGGGCAAAAGTCTTGATGATCAAGAACCACTAACAGATGATCCAGCAGCAGGACTGAG 4316  
OY 2043 GGTGCTGGGGCAAGGCGCATCTCCATGCGCATCACCTCACAAGGCCCTTGATGCTTGA 2102  
DB 4317 GGTGCTGGGGCAAGGCGCATCTCCATGCGCATCACCTCACAAGGCCCTTGATGCTTGA 4376  
OY 2103 CCATGAGGGGCGAGGAAGGTGTCTCTGGAACATGATGCGGTCTTACTTACTCT 2162  
DB 4377 CCATGAGGGGCGAGG-AGGTGTCTCTGGAACATGATGCGGTCTTACTTACTCT 4435  
OY 2163 TCTGTCCCGGCAACTGTCTCCAGATCTGTCACTATCTGAGGGGTCCNTAAGCGGG 2222  
DB 4436 TCTGTCCCGGCAACTGTCTCCAGATCTGTCACTATCTGAGGGGTCC-FAAGACAGG 4493  
OY 2223 CAGTCACTAGATCTTTTCCAGCGCATTAAGTATGAATGCGGGAGCTTTATCTTTTC 2282  
DB 4494 CAGTCACTAGATC-TTCTCCCGGCACTTAAGTATG-ACTGGGGAGCTTTATCTTTTC 4551  
OY 2283 AATGCTTTTCTAATATGCTTAAAGGCCCATCTACTTGTGAGGAGACATTTCTAG 2342  
DB 4552 AATGCTTTTCTAATATGCTTAAAGGCCCATCTACTTGTGAGGAGACATTTCTAG 4611  
OY 2343 AAAAGCAGGGGCGCATTAACCTGAACATAGGAAGAAACACCGTGTGTGTCGTC 2402  
DB 4612 AAAAGCAGGGGCGCATTAACCTGAACATAGGAAGAAACACCGTGTGTGTC-CCCC 4670  
OY 2403 TCTGTGAGGAAGAAATTAATCTGAAAGTCTGGGCAACAGAAAGCAATAGGACAGGCA 2462  
DB 4671 TCTGTGAGGAAGAAATTAATCTGAAAGTCTGGGCAACAGAAAGCAATAGGACAGG-CA 4729  
OY 2463 AAGAAAGCCGCTCTGTCAAGTTAACTAAAGATTCACCTTCCCTTCCCAAGG 2522  
DB 4730 AAGAAAGCCGCTCTGTCAAGTTAACTAAAGATTCACCTTCCCTTCCCAAGG 4789  
OY 2523 CAGTACCCCTCAGACCCCAAGGCCCAACAAGATTCCTAAAGATTTGTAAGGACTTAA 2582  
DB 4790 CAGTACCCCTCAGACCCCAAGGCCCAACAAGATTTGTAAGGACTTAA 4849  
OY 2583 GCCCAAGGCTTAAATTAATCATGATTAATCTCTGCGATTAATCCGTAAGTGAAG 2642  
DB 4850 GCCCAAGGCTTAAATTAATCATGATTAATCTCTGCGATTAATCCGTAAGTGAAG 4904  
OY 2643 GGGACAGAAACCAAGTGAAGTGAAGGTTAGTGAAGTCTCAGGATTAATCATGAG 2702  
DB 4905 ATTACAGAAACCAACAAGTGAAG-GGTTAGTGAAGTCTCAGGATTAATCAT- GAG 4962  
OY 2703 GCGGTGTCTTTTATACCAAGCTGTACCTAGCCCTTATACCTGTATTTCCCAATAC 2762  
DB 4963 GCGGTGTCTTTTATACCAAGCTGTACCTAGCCCTTATACCTGTATTTCCCAATAC 5022  
OY 2763 GAGGAAGCAGAGTGTTAACATCTCTGAGACCTTAAAGATGCTTCTTCTGATCCCTGA 2822  
DB 5023 GAGGAAGCAGAGTGTTAACATCTCTGAGACCTTAAAGATGCTTCTTCTGATCCCTGA 5082  
OY 2823 CAGTCTGACTCTAATCTTGTGTGCTTGAAGATCTTCAAAACCCATCTCACTC 2882  
DB 5083 CAGTCTGACTCTAATCTTGTGTGCTTGAAGATCTTCAAAACCCATCTCACTC 5142  
OY 2883 ACCTGAGCTTTTAAACCCCAAGGTTCAAGGATAGGCCCATCTAATTTGGCAGGACATTA 2942  
DB 5143 ACCTGAGCTTTTAAACCCCAAGGTTCAAGGATAGGCCCATCTAATTTGGCAGGACATTA 5202

OY 2943 GCCCAAGCTTGAAGCTATMTACATCTTGAACATCTTGTCTTCRGTAKGTGATGAT 3002  
DB 5203 GCCCAAGCTTGAAGCTATMTACATCTTGAACATCTTGTCTTCRGTAGGTGATGAT 5262  
OY 3003 TTAATTTTGGGCTTCAAGAACTTTGTCATCAAGCCACCCAGACGCTTTMAT 3062  
DB 5263 TTAATTTTGGGCTTCAAGAACTTTGTCATCAAGCCACCCAGACGCTTTMAT 5322  
OY 3063 TTCTCGGATCTGAGGCTTCAAGGTTTCAAGGATGCTCACTGCTCAGCAGAG 3122  
DB 5323 TTCTCGGATCTGAGGCTTCAAGGTTTCAAGGATGCTCACTGCTCAGCAGAG 5382  
OY 3123 GTTAATTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3182  
DB 5383 GT- -TACTTAAGGCTTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5438  
OY 3183 CAGCTTATAGGCTTATCTCATCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3242  
DB 5439 CAGCTTATAGGCTTATCTCATCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5498  
OY 3243 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3302  
DB 5499 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5557  
OY 3303 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3362  
DB 5558 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5617  
OY 3363 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3422  
DB 5618 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5663  
OY 3422 CCAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3482  
DB 5664 CCAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5722  
OY 3483 CCAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3542  
DB 5723 CCAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5782  
OY 3543 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3602  
DB 5783 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5842  
OY 3603 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3662  
DB 5843 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5902  
OY 3663 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3722  
DB 5903 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 5962  
OY 3723 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3782  
DB 5963 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 6022  
OY 3783 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3842  
DB 6023 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 6082  
OY 3843 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3902  
DB 6083 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 6142  
OY 3903 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 3962  
DB 6142 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 6201  
OY 3963 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 4022  
DB 6202 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 6261  
OY 4023 TTAATTAAGGCTTCAAGGATGCTCAAGGATGCTCAAGGATGCTCACTGCTCAGCAGAG 4082

Db	6262	TACGAGACCTAAAGCCCTCACCCTGGGCACTGAATTTAGAGAAAGAAAAGGGCAATA	6321
Oy	4083	TATATACAGCTCTARATATGCTTACCTAGTCNTCCATGCCCCATGMRCAATATGSAAG	4142
Db	6322	TATATACAGCTCTAAATATGCTTACTAGTCCTCATGCTCCATGCCATTCACCAATATGAAAG	6381
Oy	4143	AAAAGGAATTCCTAACTTCYAGAGAAACAACATCAATCACTAGAACCCATTAGARAT	4202
Db	6382	AAAAGGAATTCCTAACTTCYAGAGAAACAACATCAATCACTAGAACCCATTAGARAT	6441
Oy	4203	TATTAATGCGWGTACGAAACCTTAAAGAGTGGMAGCTTACACTGCGCGGGCTCATANA	4262
Db	6442	TATTAATGCGCTGTACAGAACTTAAAGAGTGGCACTTACACTGCGCGGGCTCATCAGA	6501
Oy	4263	AAGGAAAGPAAAGGAAATASAGGAATGCGCAAGCAKATATGGAACMAAAGACTG	4322
Db	6502	AAGGAAAGGAAAGGGAATAGAAAGAACTGCCCAAGATCTGAAGCCAAAGACTG	6561
Oy	4323	CAAGGAGGACCCCTCCATTAGAAATGCTTAACTTCCTCTAGATATAGGCTATCCCT	4382
Db	6562	CAAGGAGGACCCCTCCATTAGAAATGCTTAACTTCCTCTAGATATAGGCTATCCCT	6621
Oy	4383	TCCGGGAAAACCAAGCCCCGTACTACAGAGAGAAAGAAATGGGGAACCTCACAGAG-C	4441
Db	6622	TCCGGGAAACCAAGCCCCGTACTACAGAGAGAAAGAAATGGGGAACCTCACAGAGAC	6681
Oy	4442	AGTTTCTCCCTCGGGACGGTTAGCCACTGAAAGGGAATACTTTTGCCTGCACACT	4501
Db	6682	AGTTTCTCCCTCGGGATGGCTAGCCACTGAAGAGGGAATACTTTTGCCTGCACACT	6741
Oy	4502	ATCCAAATGGAATTACTTAAACCTTATCAAACTTTCACTTAAGCATGATGACACC	4561
Db	6742	AACCAATGGAATTACTTAAACCTTATCAAACTTTCACTTAAGCATGATGACACC	6801
Oy	4562	CATCARATGGCCCAATCACTTAATTAATGAGACAGGCCCTTTCAAACTATCAAGCAATTA	4621
Db	6802	CATCARATGGCCCAATCACTTAATTAATGAGACAGGCCCTTTCAAACTATTAAGCAATTA	6861
Oy	4622	KTCAGGGCCTGTGAATGTGCAARAPAAATATCCCTGCTTATCCCAAGCTCCTTCA	4681
Db	6862	GTCAAGGCTGTGAATGTGCAAGAAATATCCCTGCTTATCCCAAGCTCCTTCA	6921
Oy	4682	GGARAPCAAPAPACAGGCCATTACCTCTGAPAPAPACTGGCACTGATTTTACCACAGC	4741
Db	6922	GGAGAPCAAPAPACAGGCCATTACCTCTGAPAPAPACTGGCACTGATTTTACCACAGC	6981
Oy	4742	CCAAACCTCAGGAGATTTCAATATCTAATGCTGGGTATATCTTTACCGGGTTGGGCA	4801
Db	6982	CCAAATCTCAGGAGATTTCAATATCTAATGCTGGGTATATCTTTACCGGGTTGGGCA	7041
Oy	4802	AGGCTTCCCTGTGAGCAGAAAAAGGCCAAGAGGTAAATAAGGCACTGATTCATGAAA	4861
Db	7042	AGGCTTCCCTGTGAGCAGAAAAAGGCCAAGAGGTAAATAAGGCACTGATTCATGAAA	7101
Oy	4862	TAAATTCACAGATTCGCACTTCCCGAGGCTTACAGATGACAAATAGCCCTCTTCCAGG	4921
Db	7102	TAAATTCACAGATTCGCACTTCCCGAGGCTTACAGATGACAAATAGCCCTCTTCAAAG	7161
Oy	4922	CCACAGTAAACCAGGAGATTCACAGGCGTTAGGTATACATATCACTTCACTGCGCCT	7221
Db	7222	GAAGGCCCACTCCCTCAGGGAAGGTGAGAAAAGAAATGAAACCTCAAGGCACTCTAA	7281
Oy	5042	AAAAAGCAAAACCCAGAAACCCACTCAATGCGCTGTCTGTTGCTTATGACTTAAATA	5101
Db	7282	AAAAAGCAAAACCCAGAAACCCACTCAATGCGCTGTCTGTTGCTTATGACTTAAATA	7341
Oy	5102	GAATCTGCACTTTTCCCAAAAAGCAGACTTAAAGCCATACGAATGCTGTATGAAAGC	5161

Db	7342	GAATCTGCACTTTTCCCAAAAAGCAGCACTTAGCCATACGAAATGCTGATAGGAAGC	7401
Qy	5162	CCCTCATTAACCAATGACCTTGCTGTGAAGCCAGACAGCCCACTTATGTTGAGACATCAC	5221
Db	7402	CTTTCATTAACCAATGACCTTGCTGTGA - CCAAGACAGCCCACTTATGTTGAGACATCAC	7466
Qy	5222	CTCTTACGCCAATATTCACAAAGTTCTTAAAAATATACAGGAACCTATCCCTGAGAGA	5281
Db	7461	CTCTTACGCCAATATTCACAAAGTTCTTAAAAATATACAGGAACCTATCCCTGAGAGA	7520
Qy	5282	GGGAAAAAACAATTCACCCCMGAGCATGATATTTGTCAAGTCCCTCTCTAATTC	5341
Db	7521	GGGAAAAAACAATTCACCCCTTGATGATGATATTTGTCAAGTCCCTCTCTAATTC	7580
Qy	5342	CCCATCCCTAGATACATCTCGGGAAGACCCCTACCCAGTCATTTTATATYACCCCACTGC	5401
Db	7581	CCCATCCCTAGATACATCTCGGGAAGACCCCTACCCAGTCATTTTATCTACCCCACTGC	76410
Qy	5402	GTTTAAATGGCTGGAAGTGGAGTCTTGGAATACATCACTTGAAGTCAATCTGGAATCT	5461
Db	7641	GGTTAAATGGCTGGAAGTGGAGTCTTGGAATACATCACTGAGTCAAACTCTGGAATCT	7700
Qy	5462	GCAAAAGGAACCTGAAAATCCAGAGACAAAGCTAGCTATTCCTGTAACCTCTAGAGGA	5521
Db	7701	GCAAAAGGAACCTGAAAATCCAGAGACAAAGCTAGCTATTCCTGTAACCTCTAGAGGA	7760
Qy	5522	TTTGCGCTGCTCTTCAACAACAACAGAGAGAAATTAATTAATCATTAATCCCCCA	5581
Db	7761	TTTGCGCTGCTCTTCAACAACAACAAGAGAGAAATTAATTAATCATTAAT - ACCCA	7819
Qy	5582	TGSGCTTCCTTATCATATTTTTCTCTKTAATGTTTAACTCTTTGACCTCTCACTG	5641
Db	7820	TGGCTTCCTTATCATATTTTTCTCTTATCTGTTCTTAACTCTTCACTCTCACTG	7879
Qy	5642	CACCCCTCCATGCGCGCTGATAGCAAGTACGTCCTCCCTTACMAAGATTTCTATAGAGAA	5701
Db	7880	CACCCCTCCATGCGCGCTGATAGCAAGTACGTCCTCCCTTACMAAGATTTCTATAGAGAA	7939
Qy	5702	TGCAGCTTCGCGAATAATGATGAGCCCATCTGATAGAGATCTTTSATAGAGAAACCCCA	5761
Db	7940	TGCAGCTTCGCGAATAATGATGAGCCCATCTGATAGAGATTTTTCATAGAGAAACCCCA	7939
Qy	5762	CTTTCAGTGCACACCCCATATGCGCGCACTGCTATCACTCTGCACTCTTTCAGTGC	5821
Db	8000	CTTTCAGTGCACACCCCATATGCGCGCACTGCTATCACTCTGCACTCTTTCAGTGC	8059
Qy	5822	ATGCAAACTACTCATATATGAGACAGGAAAAATGATTAATCTAGTTGTCCTGAGAGACTTG	5881
Db	8060	ATGCAAACTACTCATATATGAGACAGGAAAAATGATTAATCTAGTTGTCCTGAGAGACTTG	8119
Qy	5882	GAGTCACTGCTGTGGAAGCTTAATCAACCCAACTGATATGTCGATGAGGAGGAGATTC	5941
Db	8120	GAGTCACTGCTGTGGAAGCTTAATCAACCCAACTGATATGTCGATGAGGAGGAGATTC	8179
Qy	5942	AAGATCAGGACAGGAAAAACATGTAAAAAGAAATTAATCTCCCACTCACCGGAGTACATG	6001
Db	8180	AAGATCAGGACAGGAAAAACATGTAAAAAGAAATTAATCTCCCACTCACCGGAGTACATA	8239
Qy	6002	GCACCTTACGCCCTTACAAAGACATAGATCTCTCAAACTACATGAAACCTCTCGTATCC	6061
Db	8240	GCGCTCTTACGCCCTTACAAAGACATAGATCTCTCAAACTACATGAAACCTCTCGTATCC	8299
Qy	6062	ATACTGCGCGGTAAAGCTATTTAAATACACCCCTCACTGAGGCTCAATGAGTCTCGGCC	6121
Db	8300	ATACTGCGCGGTAAAGCTATTTAAATACACCCCTCACTGAGGCTCAATGAGTCTCGGCC	8359
Qy	6122	AAAACCTTACTAATCTGTGGATATGCTCCCTCTGATCAATTCAGAGCCATATGTTCAATCC	6181
Db	8360	AAAACCTTACTAATCTGTGGATATGCTCCCTCTGATCAATTCAGAGCCATATGTTCAATCC	8419
Qy	6182	CTGTACTGTGAACAATGGAACAATCTTGACACAGAAATTAACAACAATTCCTCGTTTATGAG	6241
Db	8420	CTGTACTGTGAACAATGGAACAATCTTGACACAGAAATTAACAACAATTCCTCGTTTATGAG	8479

Oy	6242	GACGCTGTTCGCAATGGAATAACCCATACCTCAAACTCAACCTGATGTAATAATTA	6301
Db	8480	GACCTCTTGTTCCAAATGGAATAATACCACTACCTCAAACTCAACCTGATGTAATAATTA	8539
Oy	6302	GCAATACTACAATACAAACAATCCCAATGCATCAGGTGGGTAACTCTCCACACAAA	6361
Db	8540	GCAATACTACAGACACAACAATCCCAATGCATCAGGTGGGTAACTCTCCACACAAA	8599
Oy	6362	TAGTCTGGCTACCCCTCAGAAATATTTTTTGTCTGTGGTAACTCAGCCATCTGTGTTGA	6421
Db	8600	TAGTCTGGCTACCCCTCAGAAATATTTTTTGTCTGTGGTAACTCAGCCATCTGTGTTGA	8659
Oy	6422	ATGGCTCTTCAGAAATCTATGTGCTTCCCTCATCTTAACTAGTCCCCCAATGCACTACAA	6481
Db	8660	ATGGCTCTTCAGAAATCTATGTGCTTCCCTCATCTTAACTAGTCCCCCAATGCACTACAA	8719
Oy	6482	CTGAACAAGATTTATACGTTATGTGCATATCTAAGCCCGGACAAACAAAAGTACCCATTC	6541
Db	8720	CTGAACAAGATTTATACGTTATGTGCATATCTAAGCCCGGACAAACAAAAGTACCCATTC	8779
Oy	6542	TTCCCTTTGTTATAGGACAGAGAGTGCATAGGTCATAGGTGATGCTGGCATTTGCGGTATCA	6601
Db	8780	TTCCCTTTGTTATAGGACAGAGAGTGCATAGGTCATAGGTGATGCTGGCATTTGCGGTATCA	8839
Oy	6602	CAACCTCTACATCAGTTCTACTACAAACTATCTCAAGAACTAATATGGGACATGGAAACGG	6661
Db	8840	CAACCTCTACATCAGTTCTACTACAAACTATCTCAAGAACTAATATGGGACATGGAAACGG	8899
Oy	6662	TGCGCGATCTCCCTGGTCACTCTGGACATCAACTTAATCCCTGACAGCACTATCCCTTC	6721
Db	8900	TGCGCGATCTCCCTGGTCACTCTGGACATCAACTTAATCCCTGACAGCACTATCCCTTC	8959
Oy	6722	RAAATCGAAGAGCTTTAGACTYGCATACCGCTGAPAGAGGGGAAACCTGTTATTTTAC	6781
Db	8960	RAAATCGAAGAGCTTTAGACTYGCATACCGCTGAPAGAGGGGAAACCTGTTATTTTAC	9019
Oy	6782	GGGAAGAAATGCTGTTATTAATGTTAATCAATCCGGAATGTCATCTGAGAAATTTAAATAA	6841
Db	9020	GGGAAGAAATGCTGTTATTAATGTTAATCAATCCGGAATGTCATCTGAGAAATTTAAATAA	9079
Oy	6842	TTCSAGATGGAATTAACAAGTAAGAAGAGAGACCTCGAAACACTGGAACCTCGGGGACCTCC	6901
Db	9080	TTCSAGATGGAATTAACAAGTAAGAAGAGAGACCTCGAAACACTGGAACCTCGGGGACCTCC	9139
Oy	6902	TCAGCCRAATGATGCCCTCGGATCTTCCCTCTTAGAGACCTCTAGCAGCTAATAATTTGC	6961
Db	9140	TCAGCCRAATGATGCCCTCGGATCTTCCCTCTTAGAGACCTCTAGCAGCTAATAATTTGC	9199
Oy	6962	TACTCTCTTTTGAGACCCCTGATCTTTBACCTCCTGTTAACTGTGCTTCCAGAAATCG	7021
Db	9200	TACTCTCTTTTGAGACCCCTGATCTTTBACCTCCTGTTAACTGTGCTTCCAGAAATCG	9259
Oy	7022	AAGCTGTAAACTACAAATAGAGCCCAAGATCAGTCCAAAGACTAAGATCTACCGCAGAC	7081
Db	9260	AAGCTGTAAACTACAAATAGAGCCCAAGATCAGTCCAAAGACTAAGATCTACCGCAGAC	9319
Oy	7082	CCCTGAGACCGGCTGTATAGCCCAAGATCTGAAGTTTAAGAAGCTCAAAAGCAACCCCTCTG	7141
Db	9320	CCCTGAGACCGGCTGTATAGCCCAAGATCTGAAGTTTAAGAAGCTCAAAAGCAACCCCTCTG	9379
Oy	7142	AGGAATCTCAGTGCACAACCTCTACTACGCCCCCAATTCAAGAGAAAGACTTATAGACG	7201
Db	9380	AGGAATCTTAACTGACAAACCTCTACTACGCCCCCAATTCAAGAGAAAGACTTATAGACG	9439
Oy	7202	GTSGTGAGCCAACTTCCCAAGACACTTAAGTTTCTGTGTAAGTGGGGACATGAAAG	7261
Db	9440	GTCGTGAGCCAACTTCCCAAGACACTTAAGTTTCTGTGTAAGTGGGGACATGAAAG	9499
Oy	7262	ACAGGACTAGCTGGATTTCTTAGGCGATTAAGAATCCCTAAGCCTAAGTGGGAAGGTGA	7321
Db	9500	ACAGGACTAGCTGGATTTCTTAGGCGCTAAGAATCCCTAAGCCTAAGTGGGAAGGTGA	9559

OY		7322	CCACATCAGACTTTAAACAACGGGCGCTTGCACTTAGTGTACACCTGCACCATTACAGAGAC	7381
Dd		9660	CCACATCAGACTTTAAACAACGGGCGCTTGCACTTAGTGTACACCTGCACCATTACAGAGAC	9619
OY		7382	TCACTAATAATGCTAATTAATGAGCAAAGA CAGAGGTAAGAAATATGCCAATCATATTATTGCM	7441
Dd		9620	TCACTAATAATGCTAATTAATGAGCAAAAA CAGAGGTAAGAAATATGCCAATCATATTATTGCG	9679
OY		7442	TGAAGACACGACGAGAGGGA CAAATGATTCGGATATATAACCAAGTYTTGAGCGCGCAAC	7501
Dd		9680	TGAAGACACGACGAGAGGGA CAAATGATTCGGATATATAACCAAGTYTTGAGCGCGCAAC	9739
OY		7502	GGCAACCCCTTTAGGATCGCCCTCCCTTTGTATGAGGAGCTCTGTITTCATGCTAATTTCACT	7561
Dd		9740	GGCAACCCCTTTAGGATCGCCCTCCCTTTGTATGAGGAGCTCTGTITTCATGCTAATTTCACT	9799
OY		7562	CTATTAAATCTTGCAATGCR 7582	
Dd		9800	CTATTAAATCTTGCAATGCA 9820	
<hr/>				
RESULT 15				
LOCUS	A1101587	10229 bp	DNA	linear PRI 11-FEB-2004
DEFINITION	Pan troglodytes isolate 1 endogenous retrovirus ERV-W, ERMWEL locus, allele B, complete sequence.			
ACCESSION	A1101587			
VERSION	A1101587.1	GI:37544409		
KEYWORDS				
SOURCE				
ORGANISM	Pan troglodytes (chimpanzee)			
REFERENCE	Pan troglodytes Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan. 1 (bases 1 to 10229) Mallet,F., Bouton,O., Prudhomme,S., Cheynet,V., Oriol,G., Bonnard,B., Lucotte,G., Duret,L. and Mandrand,B. The endogenous retroviral locus ERMWEL is a bona fide gene involved in hominoid placental physiology Proc. Natl. Acad. Sci. U.S.A. 101 (6), 1731-1736 (2004)			
AUTHORS	Mallet,F., Bouton,O. and Oriol,G.			
JOURNAL	2 (bases 1 to 10229)			
PUBMED	14757826			
REFERENCE	Mallet,F., Bouton,O. and Oriol,G.			
TITLE	Direct Submission			
JOURNAL	Submitted (06-MAY-2002) Retrovirology Department, UMR 2142 CNRS-BioMerieux, Ecole Normale Supérieure de Lyon - 46 allée d'Italie, Lyon 69364 Cedex 07, France			
<hr/>				
FEATURES				
source	Location/Qualifiers			
	1..10229			
	/organism="Pan troglodytes"			
	/mol_type="genomic DNA"			
	/isolate="1"			
	/isolation_source="PBMC of chimpanzee 1"			
	/db_xref="taxon:9598"			
	/cell_type="PBMC"			
	/note="isolated by PCR; ERMWEL locus, allele B			
	endogenous_virus: ERV-W"			
	1..779			
LTR	join(248..810,2889..9820)			
mrna	/note="putative mRNA transcrip 1; gag-pol pseudogene"			
mrna	join(248..810,7592..9820)			
	/product="envelope glycoprotein"			
	/note="putative mRNA transcrip 2"			
mrna	join(248..810,9255..9820)			
	/note="putative mRNA transcrip 3; no coding region"			
CDS	7819..9435			
	/note="syncytin"			
	/codon_start=1			
	/product="envelope glycoprotein"			
	/protein_id="AAF68166.1"			
	/db_xref="GI:37544410"			
	/translation="MALPYHFLFTVLIPSTLTAPPRCRCMTSSSPFUELRNORR GNIDPGRSRFSKSTPTFTATHTPRNCISBATICMANHTMYTKIINPCPGSLGIVT ICVMYTFHTMSDGGGVODARKKHVEVISQLRVHSARPKGLDSLQHTTLAT			



HRLVSLFNTLTGLHEVSAQNPTNCWICLPLNERPVYSIPVEQNNFSTEINTTSV  
 LVBPVLSNLEIHTHSNLTCKVFSNTTDTTDSOCIRMTYPTQIYCLPSGIFPQGTGA  
 YRCLNGSESMCEFLVSPMTIYTBODLNIYVSKRNRLVPLIPVIGAGULGAG  
 TGIIGITTSOFYKLSQELNGDMERADSLVTLQDNLNLAIVLONRALDLTIE  
 RGGTCLFIEECYVUNOSGIYVEKYEIDRIORREELIRNTGPMQLSOMPWILP  
 FLRPLAIIILLFLRPTFNULVNFVSRLEAVLQMEPKOSKTKLYRPLDRPASP  
 RSDVNDIKGTPEEILTAOPLLRPNAGSS"  
 9494.10229

LTR  
 ORIGIN  
 Query Match 83.7%; Score 6344.8; DB 9; Length 10229;  
 Best Local Similarity 94.9%; Pred. No. 0;  
 Matches 6624; Conservative 194; Mismatches 111; Indels 52; Gaps 28;

Qy	606	GATGGGAAACGTTCCCGGCAAGACAAAAGCCCTTAAGACGATTCTGGABAAATGGGA	665
Db	2888	GATGGGAAACGTTCCCGGCAAGACAAAAGCCCTTAAGACGATTCTGGABAAATGGGA	2947
Qy	666	MCAATTGACCTCTGACACTTAAGAAAGAAAGCACTTATTTCTTGCAGTCCGCTG	725
Db	2948	CCAAATTGACCTCTGACACTTAAGAAAGAAAGCACTTATTTCTTGCAGTCCGCTG	3007
Qy	726	GCACTCTGAGGGAAGTAAATTAATTAACCAATCTTAAGCTAGACCTTTTGTAGAA	785
Db	3008	GCACTCTGAGGGAAGTAAATTAATTAACCAATCTTAAGCTAGACCTTTTGTAGAA	3067
Qy	786	AAGGCAATAGAGTGAAGTGCCTAATAGTCAAACTTTCTTTCAATTAAGACACTCA	845
Db	3066	AAGGCAATAGAGTGAAGTGCCTAATAGTCAAACTTTCTTTCAATTAAGACACTCA	3127
Qy	846	AATTATGTAAGAAAGTGTATTTATGCTTACAGAAAGCTTCAAGTCACTCCCTATC	905
Db	3128	AATTATGTAAGAAAGTGTATTTATGCTTACAGAAAGCTTCAAGTCACTCCCTATC	3187
Qy	906	CGAGATCCCC--GACTCCTTCCCGCAMTAAATTAAGACCCCTTCAACCAATGTGTC	963
Db	3188	CGAGATCCCCCTGACTCTCTCCCACTAATTAAGACCCCTTCAACCAATGTGTC	3247
Qy	964	AAAAGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTA	1023
Db	3248	AAAAGAGATAGACAAAAGGTAAACAGTGAACCAAGAGTGCCTAATTTCCCAATTA	3307
Qy	1024	GACCCCTCCCAAGCAGTGGAGAGAAAGAAATTCGCCCCAGCAGATGCTATGCTTTT	1083
Db	3308	GACCCCT--CCAAGCAGTGGAGAGAAAGAAATTCGCCCCAGCAGATGCTATGCTTTT	3366
Qy	1084	YTTCTCCAGACTTAAGCAATTAAGAAAGCACTTAAGTAAATTTCTCAGATAATCTGAT	1143
Db	3367	TCTCTCAGACTTAAGCAATTAAGAAAGCACTTAAGTAAATTTCTCAGATAATCTGAT	3426
Qy	1144	GGCTATATTTGATTTTCAAGGGTTAGGACAAATCTTTGATCTGCAATGAGAGATTA	1203
Db	3427	GGCTATATTTGATTTTCAAGGGTTAGGACAAATCTTTGATCTGCAATGAGAGATTA	3486
Qy	1204	TATGCTACTGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCACCATTAATCTGAC	1263
Db	3487	ATGCTACTGCTAAATCAGACACTAACCCCAATAGAGAAAGTGCACCATTAATCTGAC	3545
Qy	1264	CTGAGRTTGGGCAATCTCTGATCTCAGTCAAGTCAATGATNNGATGACACAGAA	1323
Db	3546	CTGAGRTTGGGCAATCTCTGATCTCAGTCAAGTCAATGATNNGATGACACAGAA	3602
Qy	1324	GGAAGAAATATGATTTCCCGCAGGCGCAGACAGCAATTCGACGATCCCATTTGGGG	1383
Db	3603	GGAAGAAATATGATTTCCCGCAGGCGCAGACAGCAATTCGACGATCCCATTTGGGG	3661
Qy	1384	ACACAGAAATATGATTTCCCGCAGGCGCAGACAGCAATTCGACGATCCCATTTGGGG	1443
Db	3662	ACACAGAAATATGATTTCCCGCAGGCGCAGACAGCAATTCGACGATCCCATTTGGGG	3719
Qy	1444	AAGGACTTAAGGAAATCTASGAAAGAAATCTTAAGATTTCTCAATGATCTCAACATA	1503
Db	3719	AAGGACTTAAGGAAATCTASGAAAGAAATCTTAAGATTTCTCAATGATCTCAACATA	3777

Qy	1504	CAGGGGAAAGGAAAGAAATCTTAATCTGCTTTCTGAGAGACTTAAGGAGCATTTAGAA	1563
Db	3778	CA-GGGAAAGGAAAGAAATCTTAATCTGCTTTCTGAGAGACTTAAGGAGCATTTAGAA	3836
Qy	1564	GGGTGCTCTCTGCTGACCTGACCTCTGAGAGGCACTTAATCTTAAGGCTAAGTTAT	1623
Db	3837	GGGTGCTCTCTGCTGACCTGACCTCTGAGAGGCACTTAATCTTAAGGCTAAGTTAT	3896
Qy	1624	CACTAGTCACTGACAGACATTAG-AAAAAATCTTAAGGCTGAGGAGGAGCA	1682
Db	3897	CACTAGTCACTGACAGACATTAGAAAAAATCTTAAGGCTGAGGAGGAGCA	3956
Qy	1683	AAACTTGAAGAACCTTATTTGAATCTTGACCACTTCGCTTTTATTAATAGATCAGAGGA	1742
Db	3957	AAACTTGAAGAACCTTATTTGAATCTTGACCACTTCGCTTTTATTAATAGATCAGAGGA	4016
Qy	1743	GCAGCGGAAACAGACAAACGGGATTAATAAAAAAGGCAACGCTTATGATGACCTCA	1802
Db	4017	GCAGCGGAAACAGACAAACGGGATTAATAAAAAAGGCAACGCTTATGATGACCTCA	4076
Qy	1803	GCAGGAGGACTTTGAGAGCTCTGAGAAAGGAAAAAGCTGGCAATTGAATGCTTAATA	1862
Db	4077	GCAGGAGGACTTTGAGAGCTCTGAGAAAGGAAAAAGCTGGCAATTGAATGCTTAATA	4136
Qy	1863	GGGCTTGGCTTCAAGTGGCGGTCTTAAGAGACATTTAAAAAAGATTGTCCAATGAAGTA	1922
Db	4137	GGGCTTGGCTTCAAGTGGCGGTCTTAAGAGACATTTAAAAAAGATTGTCCAATGAAGTA	4196
Qy	1923	AGCCGCTTCTGCTCAAGCCCTTAATTAAGGAAATCACTGAGAGGCGCATGCCCCA	1982
Db	4197	AGCCGCTTCTGCTCAAGCCCTTAATTAAGGAAATCACTGAGAGGCGCATGCCCCA	4256
Qy	1983	GGGCAAAAGTCTTTGATGACAGAGCCACTAACCAATGATCCAGCAGCAGACTGAG	2042
Db	4257	GGGCAAAAGTCTTCTGATGACAGAGCCACTAACCAATGATCCAGCAGCAGACTGAG	4316
Qy	2043	GGTGTGGGCAAGCGGCACTCCATGACATCACTTCAAGAGGCGCTGGGTATGCTTGA	2102
Db	4317	GGTGTGGGCAAGCGGCACTCCATGACATCACTTCAAGAGGCGCTGGGTATGCTTGA	4376
Qy	2103	CCATTGAGGCGCAGGAAAGTGTCTCTGAGACCTGAGTGGCTCTTAACTTAATCTCT	2162
Db	4377	CCATTGAGGCGCAGGAAAGTGTCTCTGAGACCTGAGTGGCTCTTAACTTAATCTCT	4435
Qy	2163	TCTGTCCGGAACAATGCTCTTCCAGATCTGTCACTATTCTGAGGAGGCTTAAGACGG	2222
Db	4436	TCTGTCCGGAACAATGCTCTTCCAGATCTGTCACTATTCTGAGGAGGCTTAAGACGG	4493
Qy	2223	CAGTCACTAGATTAATCTTTTCCAGACCTAAGTATGAACTGGGAGCTTATTTCTTTTC	2282
Db	4494	CAGTCACTAGATTAATCTTTTCCAGACCTAAGTATGAACTGGGAGCTTATTTCTTTTC	4551
Qy	2283	ACATGCTTTTCTAATTAATGCTTGAAGAGCCCACTACCTTTGTAAGGAGAGACATTTCTG	2342
Db	4552	ACATGCTTTTCTAATTAATGCTTGAAGAGCCCACTACCTTTGTAAGGAGAGACATTTCTG	4611
Qy	2343	AAAAGCAGGCGCATTTAATCACTGAAATTAAGAGAAAGAACACCGTGTGTGTTCCCC	2402
Db	4612	AAAAGCAGGCGCATTTAATCACTGAAATTAAGAGAAAGAACACCGTGTGTGTTCCCC	4670
Qy	2403	TGCTTGAAGAGAAATTAATCTTGAAGTCTGGGCAAGAAAGCAATTTGAGAGAGCA	2462
Db	4671	TGCTTGAAGAGAAATTAATCTTGAAGTCTGGGCAAGAAAGCAATTTGAGAGAGCA	4729
Qy	2463	AAGAAATCCGCTCTGTTCAAGTTAACTTAAGAAATCTTCTTCCCTTCAACAAAG	2522
Db	4730	AAGAAATCCGCTCTGTTCAAGTTAACTTAAGAAATCTTCTTCCCTTCAACAAAG	4789
Qy	2523	CAGTACCCCTCAGACCCCAAGGCCCAAGAGATTCAAAAAGATTGTTAAGAGCTTA	2582
Db	4790	CAGTACCCCTCAGACCCCAAGGCCCAAGAGATTCAAAAAGATTGTTAAGAGCTTA	4849

2583 GCCCAAGCTTAAACCATGATTAACCTCCCTGACGTAATTCGGTAGATGATGAGG 2642  
Db 4850 GCCCAAGGCTTAAACCATGATTAACCTCCCTGACGTAATTCGGTAGATGAGG 4804  
Qy 2643 GGCAAGAAACCCAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAATGAG 2702  
Db 4905 ATTACGAAACCCAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAATGAG 4962  
Qy 2703 GCCGTGTCTTTTAAACCATGATTAACCTCCCTGACGTAATTCGGTAGATGAG 2762  
Db 4963 GCTGTGTCTCTTAATGACGATGATGATGATGATGATGATGATGATGATGAT 5022  
Qy 2763 GAGGAAGCAGAGGTTTAAACCATGATTAACCTCCCTGACGTAATTCGGTAGAT 2822  
Db 5023 GAGGAAGCAGAGGTTTAAACCATGATTAACCTCCCTGACGTAATTCGGTAGAT 5082  
Qy 2823 CATCTGACTCTCAATCTGTTTGTGCTTGAAGATCTCAACCCAGATCTCAATC 2882  
Db 5083 CATCTGACTCTCAATCTGTTTGTGCTTGAAGATCTCAACCCAGATCTCAATC 5142  
Qy 2883 ACCTGACTCTTAAACCATGATTAACCTCCCTGACGTAATTCGGTAGATGAG 2942  
Db 5143 ACTGACTCTTAAACCATGATTAACCTCCCTGACGTAATTCGGTAGATGAG 5202  
Qy 2943 GCCCAAGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAATGAG 3002  
Db 5203 GCCCAAGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAATGAG 5262  
Qy 3003 TTACTTTTGGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3062  
Db 5263 TTACTTTTGGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5322  
Qy 3063 TTCTGCTGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3122  
Db 5323 TTCTGCTGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5382  
Qy 3123 GTTAAATCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3182  
Db 5383 GTTAAATCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5438  
Qy 3183 CAGCTTATCTGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3242  
Db 5439 CAGCTTATCTGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5498  
Qy 3243 RTAAAGGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3302  
Db 5499 RTAAAGGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5557  
Qy 3303 TAACTTAAATGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3362  
Db 5558 TAACTTAAATGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5617  
Qy 3363 MRAGTGGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3422  
Db 5618 MRAGTGGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5663  
Qy 3423 CCAACGAGGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3482  
Db 5664 CCAACGAGGCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5722  
Qy 3483 CCTTACAGTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3542  
Db 5723 CCTTACAGTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5782  
Qy 3543 TGTAGTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3602  
Db 5783 TGTAGTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5842  
Qy 3603 ATCTGAAGTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3662  
Db 5843 ATCTGAAGTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 5902  
Qy 3663 GAATGCAATCTCACTGCTAAAGGATCTGCTGCTGCTGCTGCTGCTGCTGCT 3722

5903 GAATGCAATCTCACTGCTAAAGGATCTGCTGCTGCTGCTGCTGCTGCTGCT 5962  
Qy 3723 TGAAGCTTATTAATCTTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTTAA 3782  
Db 5963 TGAAGCTTATTAATCTTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTTAA 6022  
Qy 3783 AGTCAATCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3842  
Db 6023 AGTCAATCTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 6082  
Qy 3843 TCAAACTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 3902  
Db 6083 TCAAACTTGAAGTGAAGTGAAGGTTAGTGAAGATCTCAGGATTAAT 6142  
Qy 3903 ACTTGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3962  
Db 6142 ACTTGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6202  
Qy 3963 TGTGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 4022  
Db 6202 TGTGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 4082  
Qy 4023 TGTGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 4142  
Db 6262 TGTGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 6321  
Qy 4083 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4142  
Db 6322 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6381  
Qy 4143 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4202  
Db 6382 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6441  
Qy 4203 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4262  
Db 6442 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6501  
Qy 4263 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4322  
Db 6502 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6561  
Qy 4323 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4382  
Db 6562 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6621  
Qy 4383 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4441  
Db 6622 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6681  
Qy 4442 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4501  
Db 6682 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6741  
Qy 4502 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4561  
Db 6742 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6801  
Qy 4562 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4621  
Db 6802 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6861  
Qy 4622 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4681  
Db 6862 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6921  
Qy 4682 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4741  
Db 6922 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 6981  
Qy 4742 TATATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 4801

Dh 6982 CCAATCTCAGGAGTTTCAAGTATCTACTAGTCTGGGTAGATATCTTTCACGGGTTGGCAG 7041  
Qy 4802 AGGCTTTCCTCTGTAGAGACAGAAAAAGGCCCAAGAGTAAATPAAGSCATAGTTCAATGAA 4861  
Db 7042 AGGCTTTCCTCTGTAGAGACAGAAAAAGGCCCAAGAGTAAATPAAGSCATAGTTCAATGAA 7101  
Qy 4862 TAAATTCAGATTCGGAATTCCTCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAGG 4921  
Db 7102 TAAATTCAGATTCGGAATTCCTCCGAGGCTTACAGATGACAAATAGCCCTGCTTCCAGG 7161  
Qy 4922 CCAAGTAAACCCAGGAGTATCCAGAGGCTTATGATATCATATCACTTACATGCGCCT 4981  
Db 7162 CCAAGTAAACCCAGGAGTATCCAGAGGCTTATGATATCATATCACTTACATGCGCCT 7221  
Qy 4982 GAAAGCCACAGTCTCAGGAGAGTGGAGAAAAATGAATGAATAAATCACTAAAGGACATCTAA 5041  
Db 7222 GAAAGCCACAGTCTCAGGAGAGTGGAGAAAAATGAATGAATAAATCACTAAAGGACATCTAA 7281  
Qy 5042 AAAAGCAAAACCCAGGAAACCCACCTCAATAGGCTGTCTGTGCTTATGACCTTAAATA 5101  
Db 7282 AAAAGCAAAACCCAGGAAACCCACCTCAATAGGCTGTCTGTGCTTATGACCTTAAATA 7341  
Qy 5102 GAATCGCAACTTCCCGCAAAAGAGAGTATGAGCCCATACGAAATGCTGTATGAAAGGC 5161  
Db 7342 GAATCGCAACTTCCCGCAAAAGAGAGTATGAGCCCATACGAAATGCTGTATGAAAGGC 7401  
Qy 5162 CCTTCATAACCAATGACCTTGTGCTTGAACCAAGACAGCCCACTTATGTTGACAGATCAC 5221  
Db 7402 CCTTCATAACCAATGACCTTGTGCTTGAACCAAGACAGCCCACTTATGTTGACAGATCAC 7460  
Qy 5222 CTCTTATGCCAAATATCAACAAGTTCTTAAACATTAACAAGAACTTATCCCTGAGAGA 5281  
Db 7461 CTCTTATGCCAAATATCAACAAGTTCTTAAACATTAACAAGAACTTATCCCTGAGAGA 7520  
Qy 5282 GGGAAAGAACTATTCACACCCWGTGACATGTATTAATGTCAGTCCCTTCCTCTAATTC 5341  
Db 7521 GGGAAAGAACTATTCACACCCWGTGACATGTATTAATGTCAGTCCCTTCCTCTAATTC 7580  
Qy 5342 CCCATCCCTAGATACATCTCTGGGAGAGACCTTACCCAGTCAATTTATATACCCCACTGC 5401  
Db 7581 CCCATCCCTAGATACATCTCTGGGAGAGACCTTACCCAGTCAATTTATATACCCCACTGC 7640  
Qy 5402 GGTAAAGTGCCTGAGTGAAGTCTTGATATCATCACTTGAAGTCAAACTCTGATACT 5461  
Db 7641 GGTAAAGTGCCTGAGTGAAGTCTTGATATCATCACTTGAAGTCAAACTCTGATACT 7700  
Qy 5462 GCCAAAGAACTCTGAAAAATCCAGAGACAAACGTACTACTTCTGGAACCTCTAGAGA 5521  
Db 7701 GCCAAAGAACTCTGAAAAATCCAGAGACAAACGTACTACTTCTGGAACCTCTAGAGA 7760  
Qy 5522 TTTGGCGCTCTCTCAACCAACCAAGAGAAAGTAACTAAATCATTAATCCGCCA 5581  
Db 7761 TTTGGCGCTCTCTCAACCAACCAAGAGAAAGTAACTAAATCATTAAT-ACCA 7819  
Qy 5582 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCSTTTTCACTCTCATCG 5641  
Db 7820 TGGSCCTCCCTTATCATATTTTCTCTKTASTGTSTTTTACCCSTTTTCACTCTCATCG 7879  
Qy 5642 CACCCCTTCATATGCGGCTGTATGACAGTACTCCCTTACGMAAGATTTCTATGAGAA 5701  
Db 7880 CACCCCTTCATATGCGGCTGTATGACAGTACTCCCTTACGMAAGATTTCTATGAGAA 7939  
Qy 5702 TGCAGCGTCCGGAATAATGATGCCCCATCGTATAGAGTCTTTSTAAGGAAACCCCA 5761  
Db 7940 TGCAGCGTCCGGAATAATGATGCCCCATCGTATAGAGTCTTTSTAAGGAAACCCCA 7999  
Qy 5762 CTTTACATGCCCAACCCATATGCCCGCAACTGTATCACTCTGCACTCTTTCATGC 5821  
Db 8000 CTTTACATGCCCAACCCATATGCCCGCAACTGTATCACTCTGCACTCTTTCATGC 8059  
Qy 5822 ATGCAAAATCTATTTATGAGACAGAAAAATGATTAATCTTATGTTGCTGAGAGACTTG 5881  
Db 8060 ATGCAAAATCTATTTATGAGACAGAAAAATGATTAATCTTATGTTGCTGAGAGACTTG 8119

Qy 5882 GAGTCACTGCTCTGTGAGACTTATCTTACCCAACTGATATGTCATGAGGCGTGAATTC 5941  
Db 8120 GAGTCACTGCTCTGTGAGACTTATCTTACCCAACTGATATGTCATGAGGCGTGAATTC 8179  
Qy 5942 AAGATCAGGCAAGAGAAAAATGTAAAGAAATATCTCCCAACTCACCGSGGTACATG 6001  
Db 8180 AAGATCAGGCAAGAGAAAAATGTAAAGAAATATCTCCCAACTCACCGSGGTACATG 8239  
Qy 6002 GCACTCTTATGCCCTTACAAAGACATAGATCTTCAAAACTCATAGAAACCTCCGTATCC 6061  
Db 8240 GCGCTCTATGAGCCCTTACAAAGACATAGATCTTCAAAACTCATAGAAACCTCCGTATCC 8299  
Qy 6062 ATATCGCTGCTGTAAGCCATTAATTAACCACTCTCATGAGGCTCATGAGTCTCGGCC 6121  
Db 8300 ATATCTGCTGTAAGCCATTAATTAATTAACCACTCTCATGAGGCTCATGAGTCTCGGCC 8359  
Qy 6122 AAAACCTTACTATGTTGATATGCTCCCTGAACTTTCARAGCATATGTTCAATCC 6181  
Db 8360 AAAACCTTACTATGTTGATATGCTCCCTGAACTTTCARAGCATATGTTCAATCC 8419  
Qy 6182 CTGTACTGAAACAATGGAACAACCTTACGACACGAATAAACAACCTTCCGTTTATGAG 6241  
Db 8420 CTGTACTGAAACAATGGAACAACCTTACGACACGAATAAACAACCTTCCGTTTATGAG 8479  
Qy 6242 GACCTCTGTTTCAATSTGGAATTAACCATACCTCAACCTCTGTTTAAATTTA 6301  
Db 8480 GACCTCTGTTTCAATSTGGAATTAACCATACCTCAACCTCTGTTTAAATTTA 8539  
Qy 6302 GCAATATCAATACAAACAACAACCTCAATGATGAGTGGTAACTTCCCAACA 6361  
Db 8540 GCAATATCAATACAAACAACAACCTCAATGATGAGTGGTAACTTCCCAACA 8599  
Qy 6362 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGATACCTCAGCTTATGTTTGA 6421  
Db 8600 TAGTCTGCTTACCTCAGGAATATTTTGTCTGTGATACCTCAGCTTATGTTTGA 8659  
Qy 6422 ATGCTCTTCAAGAACTATATGCTCTCTCAATCTTATGTCGCCCAATGRCATCTTA 6481  
Db 8660 ATGCTCTTCAAGAACTATATGCTCTCTCAATCTTATGTCGCCCAATGRCATCTTA 8719  
Qy 6482 CTGAACAAGTTTATACAGTTATATGTCATATCAACCCCGCAACAAGAGATCCATTC 6541  
Db 8720 CTGAACAAGTTTATACAGTTATATGTCATATCAACCCCGCAACAAGAGATCCATTC 8779  
Qy 6542 TTCTTTTGTATATAGAGACAGAGTCTAGGTGACTATAGTACTGACATGCGGTATCA 6601  
Db 8780 TTCTTTTGTATATAGAGACAGAGTCTAGGTGACTATAGTACTGACATGCGGTATCA 8839  
Qy 6602 CAACCTTACTCAGTTCCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 6661  
Db 8840 CAACCTTACTCAGTTCCTACTACAACTATCTCAAGAACTAAATGGGACATGGAACGGG 8899  
Qy 6662 TCGCGACTCCCTGCTCCTTGCAGAGATCAATTAATCTCCCTAGACAGATGACTTC 6721  
Db 8900 TCGCGACTCCCTGCTCCTTGCAGAGATCAATTAATCTCCCTAGACAGATGACTTC 8959  
Qy 6722 TAAATCGAAGCTTATAGACTTGTCTAACCGCTGABAGAGGGGAACTGTTATTTTATG 6781  
Db 8960 TAAATCGAAGCTTATAGACTTGTCTAACCGCTGABAGAGGGGAACTGTTATTTTATG 9019  
Qy 6782 GGGAAAGATGCTGTATTAATGTTATCAATCCGGAATGTCATCGAAGAAAGTTBAAGAA 6841  
Db 9020 GGGAAAGATGCTGTATTAATGTTATCAATCCGGAATGTCATCGAAGAAAGTTBAAGAA 9079  
Qy 6842 TTSAGATCGAATAACAAGTAKAGACAGAGACTTTCGAAACAATGAGACCTGAGGCTCC 6901  
Db 9080 TTSAGATCGAATAACAAGTAKAGACAGAGACTTTCGAAACAATGAGACCTGAGGCTCC 9139  
Qy 6902 TCAAGCAATGGAATGCTGAGATCTCCCTTATGAGACCTTATGAGACTTATATATTC 6961  
Db 9140 TCAAGCAATGGAATGCTGAGATCTCCCTTATGAGACCTTATGAGACTTATATATTC 9199



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: November 15, 2005, 21:00:58 ; Search time 1110 Seconds  
(without alignments)  
11176.802 Million cell updates/sec

Title: US-10-717-580-11

Perfect score: 7582

Sequence: 1 caccaatcggagataaacc.....tattaactgtgcartcgr 7582

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*

2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*

3: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*

4: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*

5: /cgn2\_6/prodata/1/ina/5A\_COMB.seq:\*

6: /cgn2\_6/prodata/1/ina/5B\_COMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6448	85.0	13537	US-09-949-016-17382	Sequence 17382, A
2	6124.6	80.8	8523	US-09-573-080A-21	Sequence 21, Appl
3	2994.2	33.5	77772	US-09-949-016-17417	Sequence 17417, A
4	2994.2	33.5	77797	US-09-949-016-12249	Sequence 12249, A
5	2888.6	39.1	2946	US-09-175-928-3	Sequence 3, Appl
6	2615.4	34.5	168394	US-09-949-016-13002	Sequence 13002, A
7	2220.4	29.3	145320	US-09-949-016-15858	Sequence 15858, A
8	2204	22.1	2763	US-09-949-016-5640	Sequence 5640, Ap
9	2168	28.6	8399	US-09-573-080A-26	Sequence 26, Appl
10	1837.4	24.2	2304	US-08-979-847B-87	Sequence 87, Appl
11	1507.4	19.9	276687	US-09-949-016-13840	Sequence 13840, A
12	1452.4	19.2	2364	US-08-979-847B-88	Sequence 88, Appl
13	1415.8	18.7	2391	US-08-691-563C-57	Sequence 57, Appl
14	1415.8	18.7	2391	US-09-374-766-57	Sequence 57, Appl
15	1415.8	18.7	2391	US-08-979-847B-53	Sequence 53, Appl
16	1341.8	17.7	8919	US-09-573-080A-54	Sequence 54, Appl
17	1324.2	17.5	1481	US-08-979-847B-105	Sequence 105, App
18	1312.4	17.3	1600	US-08-979-847B-207	Sequence 207, App
19	1301.8	17.2	1600	US-08-979-847B-205	Sequence 205, App
20	1282.4	16.9	1597	US-08-979-847B-205	Sequence 205, App
21	1200.2	15.8	3910	US-09-120-653D-1	Sequence 1, Appl
22	1087.2	14.3	1511	US-08-979-847B-117	Sequence 117, App
23	1069.4	14.1	1577	US-08-691-563C-89	Sequence 89, Appl
24	1069.4	14.1	1577	US-09-374-766-89	Sequence 89, Appl
25	1069.4	14.1	1577	US-08-979-847B-83	Sequence 83, Appl
26	996.8	13.1	2330	US-09-120-653D-4	Sequence 4, Appl
27	947.8	12.5	1167	US-08-691-563C-61	Sequence 61, Appl

28	947.8	12.5	1167	US-09-374-766-61	Sequence 61, Appl
29	947.8	12.5	1167	US-08-979-847B-57	Sequence 57, Appl
30	932.4	12.3	1158	US-08-471-724-1	Sequence 1, Appl
31	932.4	12.3	1158	US-08-471-969-1	Sequence 1, Appl
32	932.4	12.3	1158	US-08-384-137-1	Sequence 1, Appl
33	932.4	12.3	1158	US-08-470-006A-1	Sequence 1, Appl
34	932.4	12.3	1158	US-08-691-563C-1	Sequence 1, Appl
35	932.4	12.3	1158	US-09-200-990-1	Sequence 1, Appl
36	932.4	12.3	1158	US-09-133-411-1	Sequence 1, Appl
37	932.4	12.3	1158	US-09-374-766-1	Sequence 1, Appl
38	932.4	12.3	1158	US-08-979-847B-1	Sequence 1, Appl
39	914	12.1	99580	US-09-949-016-17411	Sequence 17411, A
40	895	11.8	68778	US-09-949-016-16406	Sequence 16406, A
41	863.4	11.4	149937	US-09-949-016-13890	Sequence 13890, A
42	784	10.3	42917	US-09-949-016-13795	Sequence 13795, A
43	722.6	9.5	1329	US-08-979-847B-108	Sequence 108, App
44	644.4	8.5	2448	US-08-691-563C-53	Sequence 53, Appl
45	644.4	8.5	2448	US-09-374-766-53	Sequence 53, Appl

ALIGNMENTS

```
RESULT 1
US-09-949-016-17382
; Sequence 17382, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 17382
; LENGTH: 13537
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17382

Query Match
Best Local Similarity 95.7%; Pred. No. 0; Length 13537;
Matches 6681; Conservative 195; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGAAACGTTCCCGCAAGACAAACGCCCCCTAAGAGCTATTCTGGAAATTGGGA 665
DB 4610 GATGGAAACGTTCCCGCAAGACAAACGCCCCCTAAGAGCTATTCTGGAAATTGGGA 4669
QY 666 MCATTGACCCCTGACACCTAAGAAAGAAAGCACTTATTTCTTCGAGCGCGCTG 725
DB 4670 CCAATTGACCCCTGACACCTAAGAAAGAAAGCACTTATTTCTTCGAGCGCGCTG 4729
QY 726 GCATCTCGAGGAGAGTAAATTATAACACCATTTACAGCTAGACTTTTGTAGAA 785
DB 4730 GCACTCCGAGGAGAGTAAATTATAACACCATTTACAGCTAGACTTTTGTAGAA 4789
QY 786 AAGCAATGAGAGTAAAGTGCATTAAGTCAAACTTTCTTTCATTAAGAGCAACTGAC 845
DB 4790 AAGCAATGAGAGTAAAGTGCATTAAGTCAAACTTTCTTTCATTAAGAGCAACTGAC 4849
QY 846 AATTATTAATAAGTGTATTGATTAAGCCCTACAGGAAGCTTACAGTCTACCTCCATC 905
DB 4850 AATTATTAATAAGTGTATTGATTAAGCCCTACAGGAAGCTTACAGTCTACCTCCATC 4909
QY 906 CCAATCCCGAGCTCTCCCTCCCAATTAAGAGACCCCTTCAACCAATGATGCA 965
```



4910 CCAGCATCCCGACTCTTCCCACTAATAGAACCCCTTCAACCCAAATGGTCCAA 4969  
966 AAGGAGATAGCAAAAGGGTAAACAGTGAACCAAGAGTGCATATTTCCCAATTATGA 1025  
4970 AAGGAGATAGCAAAAGGGTAAACAGTGAACCAAGAGTGCATATTTCCCAATTATGA 5029  
1026 CCCCTCCCAAGCAGTGGGAGGAAGAAATTCCGCCACGCGAGAGTGCATGCTTTTYY 1085  
5030 CCCCT-CCAAGCAGTGGGAGGAAGAAATTCCGCCACGCGAGAGTGCATGCTTTTTC 5088  
1086 TCTCCAGACTTAAAGCAATTAACAGCTTAAGTAAATTCTCAAGTAAATTCCTGATGG 1145  
5089 TCTCCAGACTTAAAGCAATTAACAGCTTAAGTAAATTCTCAAGTAAATTCCTGATGG 5148  
1146 CTATATGTTGTTTCAAGAGTTAGACAATCTTGTATCTGACATGAGAGATATATA 1205  
5149 CTATATGTTGTTTCAAGAGTTAGACAATCTTGTATCTGACATGAGAGATATATA 5207  
1206 TGTCACTGCTAAATCAAGACATAACCCCAATGAGAGAGTCCCACTTAAGTCCAGCT 1265  
5208 TGTCACTGCTAAATCAAGACATAACCCCAATGAGAGAGTCCCACTTAAGTCCAGCT 5267  
1266 GAAGTTTGGCGATCTCTGTTATCTCACTCAGTCAATGGAATGAGATGAACAAG 1325  
5268 GAAGTTTGGCGATCTCTGTTATCTCACTCAGTCAATGGAATGGAATGGAATGGA 5324  
1326 AAAAGAAATGATTTCCCAAGCGCCAGACAGAGTCCCACTTAAGTCCCACTTAAGT 1385  
5325 AAAAGAAATGATTTCCCAAGCGCCAGACAGAGTCCCACTTAAGTCCCACTTAAGT 5381  
1386 ACAGAAATCAGTAACATGAGAGATGGTGTGCTGACAGATTTGCTTAATTTGTCTA 1445  
5382 ACAGAAATCAGTAACATGAGAGATGGTGTGCTGACAGATTTGCTTAATTTGTCTA 5440  
1446 GGAATAAGAAATCTAGAAAGAAATCTTAATTAATTAATTAATTAATTAATTAATTA 1505  
5441 GGAATAAGAAATCTAGAAAGAAATCTTAATTAATTAATTAATTAATTAATTAATTA 5499  
1506 GGGGAAAGGAAAGAAATCTAGAAAGAAATCTTAATTAATTAATTAATTAATTAATTA 1565  
5500 -GGGAAAGGAAAGAAATCTAGAAAGAAATCTTAATTAATTAATTAATTAATTAATTA 5558  
1566 GTGCTCTCTGTCACTGACTCTTCTGAAGCCCACTTAATTTAAAGCTTAATTTAA 1625  
5559 GTGCTCTCTGTCACTGACTCTTCTGAAGCCCACTTAATTTAAAGCTTAATTTAA 5618  
1626 CTCACTGAGTGCAGACATTAAG-AAAAAATTTCAAAAGTCTGCGCTGAGCCCGGAG 1684  
5619 CTCACTGAGTGCAGACATTAAG-AAAAAATTTCAAAAGTCTGCGCTGAGCCCGGAG 5678  
1685 ACTTAGAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1744  
5679 ACTTAGAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5738  
1745 AGGCGGAACAGCAAAACGGGATTAATAAAAGGCAACGCTTTAAGTCAATGACCTTCA 1804  
5739 AGGCGGAACAGCAAAACGGGATTAATAAAAGGCAACGCTTTAAGTCAATGACCTTCA 5798  
1805 CAAGTGAATTTGAGGCTCTGGAAGGAAAGCTGGGCAATTTGAATGCTTAATGAG 1864  
5799 CAAGTGAATTTGAGGCTCTGGAAGGAAAGCTGGGCAATTTGAATGCTTAATGAG 5858  
1865 GCTTGCTTCAAGTCCGCTCAAGAGCACTTTAATAAAAGTTGCTCAAGTGAAGTAA 1924  
5859 GCTTGCTTCAAGTCCGCTCAAGAGCACTTTAATAAAAGTTGCTCAAGTGAAGTAA 5918  
1925 CCGGCTTCTGCTCATGCTCCCTTAATTTCAAGGAAATCACTGAAGGCTCACTGCTCC 1984  
5919 CCGGCTTCTGCTCATGCTCCCTTAATTTCAAGGAAATCACTGAAGGCTCACTGCTCC 5978  
1985 GGAACAAAGTCTTTGAGTCAAGAGCACTTAACAGATGATTCAGACAGAGAGCTGAGG 2044

5979 GGAACAAAGTCTTTGAGTCAAGAGCACTTAACAGATGATTCAGACAGAGAGCTGAGG 6038  
2045 TGCCTGGGCAAGCGCATCCATCCATCAAGAGCTGGGATATGCTTGAAC 2104  
6039 TGCCTGGGCAAGCGCATCCATCCATCAAGAGCTGGGATATGCTTGAAC 6098  
2105 ATTGAGGCAAGAGGTTGCTCCGAGCACTGGGAGGCTTTAAGTCACTTTC 2164  
6099 ATTGAGGCAAGAGGTTGCTCCGAGCACTGGGAGGCTTTAAGTCACTTTC 6157  
2165 TGTCCGAGCAACTGCTCTCAAGATCTGCACTATTTCTGAGGGGCTCCTTAAGCGGCA 2224  
6158 TGTCCGAGCAACTGCTCTCAAGATCTGCACTATTTCTGAGGGGCTCCTTAAGCGGCA 6215  
2225 GTCACTAGATCTTTTCCAGCCACTTAAGTGAATGAGCTGGGAGCTTTATTTCTTAC 2284  
6216 GTCACTAGATCTTTTCCAGCCACTTAAGTGAATGAGCTGGGAGCTTTATTTCTTAC 6273  
2285 ATGCTTTCTAATTAATGCTTGAAGCCCACTTAAGTGAAGAGACATTTAGCA 2344  
6274 ATGCTTTCTAATTAATGCTTGAAGCCCACTTAAGTGAAGAGACATTTAGCA 6333  
2345 AAGCAGGCGCAATTAACCTGAACATAGAGAGAAACCCGTTTGTGTTCCCTG 2404  
6334 AAGCAGGCGCAATTAACCTGAACATAGAGAGAAACCCGTTTGTGTTCCCTG 6392  
2405 CTGAGAGAGAAATTAATCTGGAAGTCTGGGCAAGAGAGCAATATGAGAGCA 2464  
6393 CTGAGAGAGAAATTAATCTGGAAGTCTGGGCAAGAGAGCAATATGAGAGCA 6451  
2465 GAATGCGCCGCTGCTCAAGTTAACTTAAGAGATTCACCTTCTTCCCTCAAGAGCA 2524  
6452 GAATGCGCCGCTGCTCAAGTTAACTTAAGAGATTCACCTTCTTCCCTCAAGAGCA 6511  
2525 GTACCCCTCAAGCCCAAGGCGCCCAAGAGATTCAAAGATTTAAGACTTAAGAGCA 2584  
6512 GTACCCCTCAAGCCCAAGGCGCCCAAGAGATTCAAAGATTTAAGAGACTTAAGAGCA 6571  
2585 CCAAGGCTTAATTAATCACTGCACTAATCTCCCTGAGTAAATTCCTGAGTGAATGAGAG 2644  
6572 CCAAGGCTTAATTAATCACTGCACTAATCTCCCTGAGTAAATTCCTGAGTGAATGAGAG 6624  
2645 CAGAGAAACCACTGAGAGTGAAGGATTAAGTCAAGATCTCAGATTAATCAATGAGAG 2704  
6625 TACAGAAACCACTGAGAGTGAAGGATTAAGTCAAGATCTCAGATTAATCAATGAGAG 6682  
2705 GCTTGTCTTTTAATCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2764  
6683 TGTGTTCTCTTAATCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6742  
2765 GGAAGCAAGTGGTTTAATCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2824  
6743 GGAAGCAAGTGGTTTAATCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6802  
2825 TCTGACTCTCAATCTGTTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2884  
6803 TCTGACTCTCAATCTGTTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6862  
2885 CTGGAATTTTCAAGGAGTGAAGGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2944  
6863 CTGGAATTTTCAAGGAGTGAAGGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 6922  
2945 CCAAGACTTAATTAATCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3004  
6923 CCAAGACTTAATTAATCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6980  
3005 ACTTTGAGCGCTTAATTAATCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3064  
6981 ACTTTGAGCGCTTAATTAATCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 7040  
3065 CTTGCGCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3124  
7041 CTTGCGCACTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 7100

QY 3125 TAAATACCTTAGRCTAARATTAATCCAAAGGACAGAGGCCCTCAGTAGGAAYATACCA 3184  
Db 7101 ----TACTTAGGGCTAAATTAATCCAAAGGACAGAGGCCCTCAGTAGGAACACATCCA 7156  
QY 3185 GGCCTTACGCGGCTTATCCCTCATCYCAAAAACCTTAAAGCACTAAGRRRTCTTGGCCT 3244  
Db 7157 GCTTATACGCGCTTATCCCTCATCCCAAAAACCTTAAAGCACTAAGGGAGTCTTGGCGCT 7216  
QY 3245 AAYAGYTTCTGCGCAAAATGAGATCCCGAGGTTMGCGCAAAATAGCGAGGYATATATAT 3304  
Db 7217 AATAGTTTCTGCCCAAAATGAGAT-CCAGGATAGCGAAATAGCCAGGTCAATTAATTA 7275  
QY 3305 CASTAATTAAGGAAATCAGAAAGCCAAATACCATTATTAAGTAGAAYAMCTGAAYMR 3364  
Db 7276 CACTAATTAAGGAAATCAGAAAGCCAAATACCATTATTAAGTAGAAYAMCTGAAYMR 7335  
QY 3365 AAGTGGCTTTCCAGGCCCTTAAAGAGCCTTAAACCCAGYCCCAAGTTTAAGYTTGCC 3424  
Db 7336 AAGTGGCTTTCCAGGCC-----TAAACCAAGCCCAAGTGTTAAGTTTGC 7381  
QY 3425 AACGGGCAAGACATTTTSTYTAAYRTCAAGAAAAAACAABAYACCTCTRGAGTCC 3484  
Db 7382 AACGGGCAAGACATTTTCTCATATGTACAG-AAAAACAAGAAATACCTCTAGAGTCC 7440  
QY 3485 TTACACAGRTCCRAGGAYGAGCTTGCAACCTGTGCGRACCTGASTAAGGAAATGATG 3544  
Db 7441 TTACACAGATCCAGGGAGTGAGCTTGCAACCTGTGCGATACCTGATAGGAAATGATG 7500  
QY 3545 TAGTGGCAAAAGGTTGRCYATTTGTTTAYGGGTAGTGTGGCAGTAGCACTYKTATAT 3604  
Db 7501 TAGTGGCAAAAGGTTGACCTCATTTGTTTACGGGTAGTGTGGCAGTAGCACTYKTATAT 7560  
QY 3605 CTGAAGCAGTTAAATATAATACAGGAGAGAGATCTTACGTGTGACATCTCATAGTGA 3664  
Db 7561 CTGAAGCAGTTAAATATAATACAGGAGAGAGATCTTACGTGTGACATCTCATAGTGA 7620  
QY 3665 AYAGCATACTCACTGTCAAAGAGACTGTGTGCTGACAAACGTGTTACTTAATRTC 3724  
Db 7621 ATGAGCATACTCACTGTCAAAGAGACTGTGTGCTGACAAACGTGTTACTTAATRTC 7680  
QY 3725 AGGCTCAATTACTTGAARGGCAGTGTGCACTGTGCACTTGTGCACTTTAAACCAAG 3784  
Db 7681 AGGCTCAATTACTTGAARGGCAGTGTGCACTGTGCACTTGTGCACTTTAAACCAAG 7740  
QY 3785 YNCATTTCTTCCAGAGCAATGAAGAAAGATTAATTAATCTGTCAACATTAATTTCTC 3844  
Db 7741 CCACATTTCTTCCAGAGCAATGAAGAAAGATTAATTAATCTGTCAACATTAATTTCTC 7800  
QY 3845 AAACCTATGCGCACTCGAGGGGACCTTGTAGARGTTCYTTGACGTGACGACCTTCAAC 3904  
Db 7801 AAACCTATGCGCACTCGAGGGGACCTTGTAGARGTTCYTTGACGTGACGACCTTCAAC 7859  
QY 3905 TTGTATACGTAGGAAAGTTCTTTGTAGAAAAAGACTTCGAAAAAGYGGGTTATGCACTG 3964  
Db 7860 TTGTATACGTAGGAAAGTTCTTTGTAGAAAAAGACTTCGAAAAAGYGGGTTATGCACTG 7919  
QY 3965 GTACAGGATTAATGGAATTAATTTGAAAGTAATCCCTCACTCGAGAACTAGTCTTAACT 4024  
Db 7920 GTACAGGATTAATGGAATTAATTTGAAAGTAATCCCTCACTCGAGAACTAGTCTTAACT 7979  
QY 4025 GCGAAGCTAATAGCCYTCATYTKGGGCACTAGAAATTAAGAGAAAGAAAGGAAATATA 4084  
Db 7980 GCGAAGCTAATAGCCYTCATYTKGGGCACTAGAAATTAAGAGAAAGAAAGGAAATATA 8039  
QY 4085 TTATACGACTCTAATATAGCTTAACTAGTCTTCAATGCTGATGCTGATGCTGATGCTGAT 4144  
Db 8040 TTATACGACTCTAATATAGCTTAACTAGTCTTCAATGCTGATGCTGATGCTGATGCTGAT 8099  
QY 4145 AGGAAATTCCTAACTTCYAGGAGAACACCTATCMAATCAGAGAACCTTAGAGAAATTA 4204  
Db 8100 AGGAAATTCCTAACTTCYAGGAGAACACCTATCMAATCAGAGAACCTTAGAGAAATTA 8159

QY 4205 TTAATGCGGATACAGAAACCTAAGAGGTGGAAGTCTTACATGCTGCGGGCTCATCMAAA 4264  
Db 8160 TTAATGCGGATACAGAAACCTAAGAGGTGGAAGTCTTACATGCTGCGGGCTCATCMAAA 8219  
QY 4265 GGAAGAAAGGAAATTAATGAAGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAAT 4324  
Db 8220 GGAAGAAAGGAAATTAATGAAGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAAT 8279  
QY 4325 AGGAGAGCCTCCATTAAGAAATGCTTAACTTCCCTTAAGTAAAGGTATATCCCTTC 4384  
Db 8280 AGGAGAGCCTCCATTAAGAAATGCTTAACTTCCCTTAAGTAAAGGTATATCCCTTC 8339  
QY 4385 GGGGAAACCAAGCCCGCATCTAGCAGAGGAAACAGAAATGGGAAACCTCAGAGG-CAG 4443  
Db 8340 GGGGAAACCAAGCCCGCATCTAGCAGAGGAAACAGAAATGGGAAACCTCAGAGGAG 8399  
QY 4444 TTTTCTCCCTCGGAGCGTTAGGCACTGAAGAGGAAATTACTTTGCGCTGCAACTAT 4503  
Db 8400 TTTTCTCCCTCGGAGCGGCTAGGCACTGAAGAGGAAATTACTTTGCGCTGCAACTAT 8459  
QY 4504 CCAATGGAATTACTTAAACCTTCATCAACCTTTCACTTAGGCAATCGATAGCAACCA 4563  
Db 8460 CCAATGGAATTACTTAAACCTTCATCAACCTTTCACTTAGGCAATCGATAGCAACCA 8519  
QY 4564 TCAATGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4623  
Db 8520 TCAATGCGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8579  
QY 4624 CAGGCGCTGTGAATGTGCGCAABAAATATCCCTGCTATATGCGCAAGCTCTTCAAG 4683  
Db 8580 CAGGCGCTGTGAATGTGCGCAABAAATATCCCTGCTATATGCGCAAGCTCTTCAAG 8639  
QY 4684 ABAACAAABAACAGGCAATTAACCTGTGABABAACCTGCAACCTGTTTACCAACAGCC 4743  
Db 8640 ABAACAAABAACAGGCAATTAACCTGTGABABAACCTGCAACCTGTTTACCAACAGCC 8699  
QY 4744 AAACCTCAGGGATTTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 4803  
Db 8700 AAACCTCAGGGATTTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8759  
QY 4804 GCGTTCCTGTGAGCAGAAAAAGGCCAAGAGGTATATAAGGCACTAGTTCATGAATA 4863  
Db 8760 GCGTTCCTGTGAGCAGAAAAAGGCCAAGAGGTATATAAGGCACTAGTTCATGAATA 8819  
QY 4864 ATTCCAGATTTGGAATTTCCCGAGGCTTACAGAGTGAACATAGCCCTGCTTCAAGGCC 4923  
Db 8820 ATTCCAGATTTGGAATTTCCCGAGGCTTACAGAGTGAACATAGCCCTGCTTCAAGGCC 8879  
QY 4924 ACAATTAACCAAGGAGTATCCCAAGGCTTATAGTATACATATCACTTACATCAGCGCTGA 4983  
Db 8880 ACAATTAACCAAGGAGTATCCCAAGGCTTATAGTATACATATCACTTACATCAGCGCTGA 8939  
QY 4984 AGGCGCAGTCTCAGGGAAGGTGAGAAATGAAGAAATGAAGAAATGAAGAAATGAAGAA 5043  
Db 8940 AGGCGCAGTCTCAGGGAAGGTGAGAAATGAAGAAATGAAGAAATGAAGAAATGAAGAA 8999  
QY 5044 AAGCAAAACCAAGAAACCAACCTCAATGCGCTGTCTGTGCTTAAAGCTTAAAGAA 5103  
Db 9000 AAGCAAAACCAAGAAACCAACCTCAATGCGCTGTCTGTGCTTAAAGCTTAAAGAA 9059  
QY 5104 ATCTGCAATTTCCCAAAAACAGAGACTTAAGCCCATAGAAATGCTGATAGAGGCC 5163  
Db 9060 ATCTGCAATTTCCCAAAAACAGAGACTTAAGCCCATAGAAATGCTGATAGAGGCC 9119  
QY 5164 TTCAATACCAATGACCTTGTGCTTGAACCAAGACAGCACTTAGTTGAGACATCACT 5223  
Db 9120 TTCAATACCAATGACCTTGTGCTTGAACCAAGACAGCACTTAGTTGAGACATCACT 9179  
QY 5224 CTTAGCCAAATATCAAGAGTTCTTAAACATTAACAGAACTATCCCTGAGAGAGG 5283  
Db 9180 CTTAGCCAAATATCAAGAGTTCTTAAACATTAACAGAACTATCCCTGAGAGAGG 9239  
QY 5284 GAAAGAACTATTCACCCMMGTGACATGATTAATGTAAGTCCCTCTCTAATTTCC 5343

Db	9240	GAAGAAAGCAATTCACCCCTTGGACATGGATATTAGCAAGTCCCTTCCTTAATTC	9299
Oy	5344	CATCCCTAGATATCATCTCTGGGAAGAGACCCCTACCCAGTCAATTTATATACCCCACTGCGG	5403
Db	9300	CATCCCTAGATATCATCTCTGGGAAGAGACCCCTACCCAGTCAATTTATATACCCCACTGCGG	9359
Oy	5404	TTAAAGTGGCTGAGTGGAGCTTTGGATATACATCACTTGAGTCAAAATCCTGAGTATG	5463
Db	9360	TTAAAGTGGCTGAGTGGAGCTTTGGATATATCATCACTTGAGTCAAAATCCTGAGTATG	9419
Oy	5464	CAAGGAACCTGAAATATCCAGGAGACAAACCTTAGCTATTCCTGTGAACCTCTAGAGGATT	5523
Db	9420	CAAGGAACCTGAAATATCCAGGAGACAAACCTTAGCTATTCCTGTGAACCTCTAGAGGATT	9479
Oy	5524	TGCGCTGCTCTTCAACAAACAACAGAGAAAGTAAGTAATAATCATAAATCCCCATG	5583
Db	9480	TGCGCTGCTCTTCAACAAACAACAGAGAAAGTAAGTAATAATCATAAATCCCCATG	9538
Oy	5584	GSCTCCCTATCATATTTTTCTCTKTATGTSTTTTAAACCTSTTTCACTCTCATGCA	5643
Db	9539	GSCTCCCTATCATATTTTTCTCTKTATGTSTTTTAAACCTSTTTCACTCTCATGCA	9598
Oy	5644	CCCCCTCATATGCCCTGTATGACCAAGTATGCCCTTACCMAGATTTCTAAGAGAAAG	5703
Db	9599	CCCCCTCATATGCCCTGTATGACCAAGTATGCCCTTACCMAGATTTCTAAGAGAAAG	9658
Oy	5704	CAGGTCGCCGAATATTTGATGCCCATTCGATAGAGTCTTSTTAAGGAAACCCCAAC	5763
Db	9659	CAGGTCGCCGAATATTTGATGCCCATTCGATAGAGTCTTSTTAAGGAAACCCCAAC	9718
Oy	5764	TTCACTGCCCAACCCCATATGCCCCCGCACTGCTATCACTCTGCACTCTTTGATGAT	5823
Db	9719	TTCACTGCCCAACCCCATATGCCCCCGCACTGCTATCACTCTGCACTCTTTGATGAT	9778
Oy	5824	GCAAAATCACTATTTGACAGGAAATATGATTAATCTTAGTTGTCCTGAGGACTTTGGA	5883
Db	9779	GCAAAATCACTATTTGACAGGAAATATGATTAATCTTAGTTGTCCTGAGGACTTTGGA	9838
Oy	5884	GTCACGTCTGTGGAATTAATTCACCCAACTGGTATGCTATGGGGGTGAGTTCA	5943
Db	9839	GTCACGTCTGTGGAATTAATTCACCCAACTGGTATGCTATGGGGGTGAGTTCA	9898
Oy	5944	GATCAGGCAAGAAACATGTAAAAAGATTAATCTCCCACTCACSSGGGTATATGTC	6003
Db	9899	GATCAGGCAAGAAACATGTAAAAAGATTAATCTCCCACTCACSSGGGTATATGTC	9958
Oy	6004	ACCTCTAGCCCCCTACAAAGACTAGATCTCAAAATCAATGAAACCTCTGTAACCT	6063
Db	9959	ACCTCTAGCCCCCTACAAAGACTAGATCTCAAAATCAATGAAACCTCTGTAACCT	10018
Oy	6064	ACTGCGCTGTGAAGCTATTTAATACCAACCTCATGCGGGCCATGAGTCTGCGGCCAA	6123
Db	10019	ACTGCGCTGTGAAGCTATTTAATACCAACCTCATGCGGGCCATGAGTCTGCGGCCAA	10078
Oy	6124	AACCTTACTAATGTGTGATATGCTCCCTCCCTGAACTTCARGCCATATGTTCAATCCCT	6183
Db	10079	AACCTTACTAATGTGTGATATGCTCCCTCCCTGAACTTCARGCCATATGTTCAATCCCT	10138
Oy	6184	GTAACCTGAACATGGAACAACCTTCAGCAGAGAAATAAACACCACTTCCGTTTTAGTAGGA	6243
Db	10139	GTAACCTGAACATGGAACAACCTTCAGCAGAGAAATAAACACCACTTCCGTTTTAGTAGGA	10198
Oy	6244	CCTCTTGTTCCTCAATSTGGAATTAACCATACCTCAACCTCACTGTGTAAATTTAGC	6303
Db	10199	CCTCTTGTTCCTCAATSTGGAATTAACCATACCTCAACCTCACTGTGTAAATTTAGC	10258
Oy	6304	AATATCATATACACAACCACTCCCAATGATCAAGTGGGTACTCTCCCAACAATA	6363
Db	10259	AATATCATATACACAACCACTCCCAATGATCAAGTGGGTACTCTCCCAACAATA	10318
Oy	6364	GTCTGCTTACCTCAGGAATATTTTTGTCTGTGATCTCAGCTTATCGTTGTTGAAT	6423

Db	10319	GTCTGCTACCTCAGGAATATTTTGTCTGTGTACTCAGCCTATCGTTGTTGAAT	10378
Qy	6424	GGCTCTTCAGATCTATGTGCTTCCCTCATTTCTTTAGTGCCTTCATCTACCT	6483
Db	10379	GGCTCTTCAGATCTATGTGCTTCCCTCATTTCTTTAGTGCCTTCATCTACCT	10438
Qy	6484	GAACAGATTTATACGTTATGTCTATCTAAAGCCCCCACAATAAGATACCCATCTT	6543
Db	10439	GAACAGATTTATACGTTATGTCTATCTAAAGCCCCCACAATAAGATACCCATCTT	10498
Qy	6544	CCTTTGTTATAGGAGCAGAGTGCTAGTGTCACAGGTACTAGGCAATTGGCGTATCACA	6603
Db	10499	CCTTTGTTATAGGAGCAGAGTGCTAGTGTCACAGGTACTAGGCAATTGGCGTATCACA	10558
Qy	6604	ACCTTACTCAGTTTCTACTACAACTATCTCAAGACTAATAGGAGCATGAAACGGGTC	6663
Db	10559	ACCTTACTCAGTTTCTACTACAACTATCTCAAGACTAATAGGAGCATGAAACGGGTC	10618
Qy	6664	GCCGACTCCCTGGTCACTTGGCAAGTCACTTACTCCCTAGCAGCAGATGCTCTCA	6723
Db	10619	GCCGACTCCCTGGTCACTTGGCAAGTCACTTACTCCCTAGCAGCAGATGCTCTCA	10678
Qy	6724	AATGAAAGCTTTAGACTGCTGTAACCGCTGABAGAGGGAGCAACGTTTATTTTAGGG	6783
Db	10679	AATGAAAGCTTTAGACTGCTGTAACCGCTGABAGAGGGAGCAACGTTTATTTTAGGG	10738
Qy	6784	GAAGATCTGTTATTTATGTTATATCAATCCGGAATGTCCTAGAGAAAGTTAAGAAAT	6843
Db	10739	GAAGATCTGTTATTTATGTTATATCAATCCGGAATGTCCTAGAGAAAGTTAAGAAAT	10798
Qy	6844	CSAGATCGAATTCAGCTTAKAGCAGABAGCTTCCAAAACATGAGCCCGGGCCCTC	6903
Db	10799	CSAGATCGAATTCAGCTTAKAGCAGABAGCTTCCAAAACATGAGCCCGGGCCCTC	10858
Qy	6904	AGCCATGGAATGCGCGGATTCCTCCCTCTTAGAGCCCTTAGCAGCTAATATTGCTA	6963
Db	10859	AGCCATGGAATGCGCGGATTCCTCCCTCTTAGAGCCCTTAGCAGCTAATATTGCTA	10918
Qy	6964	CTCCTCTTTGGACCCCTGTATCTTTTACCTCCTGTTTAACTTTGCTCTTCAGAAATCGAA	7023
Db	10919	CTCCTCTTTGGACCCCTGTATCTTTTACCTCCTGTTTAACTTTGCTCTTCAGAAATCGAA	10978
Qy	7024	GCTGTAAACCTAACAAATGAGGCCCAAGATGCACTAAGATCTAACCGCAGACC	7083
Db	10979	GCTGTAAACCTAACAAATGAGGCCCAAGATGCACTAAGATCTAACCGCAGACC	11038
Qy	7084	CTGAGACCGGCGTATAGCCCAAGATCTGATGTTAATGATCAATCAAGCACCCTCTGAG	7143
Db	11039	CTGAGACCGGCGTATAGCCCAAGATCTGATGTTAATGATCAATCAAGCACCCTCTGAG	11098
Qy	7144	GAATCTCAGCTGCACAACCTCTACTACGCCCAATTCAGAGAAAGCAATTAGCGGCT	7203
Db	11099	GAATCTCAGCTGCACAACCTCTACTACGCCCAATTCAGAGAAAGCAATTAGCGGCT	11158
Qy	7204	SGTGGGCAACCTCCCAACAGACTTTAGGTTTCTGTGTTGAGATGGGGGACTGAGAGAC	7263
Db	11159	SGTGGGCAACCTCCCAACAGACTTTAGGTTTCTGTGTTGAGATGGGGGACTGAGAGAC	11218
Qy	7264	AGGACTAGTGGATTTCTTAGGTGATGATTAAGATCTCTTAAGCCTTAGSTGGGAAGGTACCC	7323
Db	11219	AGGACTAGTGGATTTCTTAGGTGATGATTAAGATCTCTTAAGCCTTAGSTGGGAAGGTACCC	11278
Qy	7324	ACATCACCCTTTAAACAGCGGGCTTGCACTTAAGTCAACCTGACCAATCAAGAGCTC	7383
Db	11279	ACATCACCCTTTAAACAGCGGGCTTGCACTTAAGTCAACCTGACCAATCAAGAGCTC	11338
Qy	7384	ACTAAATGCTAATTAGGCAAGACGAGGTAAAGAAATAGCCAAATCAATCTATGCTTG	7443
Db	11339	ACTAAATGCTAATTAGGCAAGACGAGGTAAAGAAATAGCCAAATCAATCTATGCTTG	11398
Qy	7444	AGAGCAGACAGAGGAGCAATGATGGGATATTAACCCAAAGTCTTGAGACCGGCAACGG	7503
Db	11399	AGAGCAGACAGAGGAGCAATGATGGGATATTAACCCAAAGTCTTGAGACCGGCAACGG	11458

QY 7504 CAACCCCTTGGGTCCTTCTTGTATGAGAGCTCTGTTTCATGCTATTTCACTCT 7563  
DB 11459 CAACCCCTTGGGTCCTTCTTGTATGAGAGCTCTGTTTCATGCTATTTCACTCT 11518  
QY 7564 ATTAATCTTGCACCTGCR 7582  
DB 11519 ATTAATCTTGCACCTGCA 11537

## RESULT 2

US-09-573-080A-21  
Sequence 21, Application US/09573080A  
Patent No. 6828097  
GENERAL INFORMATION:  
APPLICANT: JOAN, KNOWL  
APPLICANT: ROGAN, PETER  
TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERATI  
FILE REFERENCE: 30307  
CURRENT APPLICATION NUMBER: US/09/573,080A  
CURRENT FILING DATE: 2000-05-16  
NUMBER OF SEQ ID NOS: 479  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 21  
LENGTH: 8523  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: repeat region  
LOCATION: (1)..(8523)  
OTHER INFORMATION: herx17  
PUBLICATION INFORMATION:  
AUTHORS: Jurka, J; Malchewicz, J; Milosavljevic, A  
TITLE: Prototypic sequences for human repetitive DNA  
JOURNAL: Journal of Molecular Evolution  
VOLUME: 35  
ISSUE: 4  
PAGES: 286-291  
DATE: 1992-10-  
DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
DATABASE ENTRY DATE: 1996-01-26  
DATABASE ENTRY DATE: 1996-01-26  
US-09-573-080A-21

Query Match 80.8%; Score 6124.6; DB 4; Length 8523;

Best Local Similarity 95.7%; Pred. No. 0; Matches 6363; Conservative 186; Mismatches 50; Indels 53; Gaps 27;

QY 606 GATGGGAAAGCTTCCCGCAAGCAAAAGCGCCCTAAGAGCTATTCGARAATTGGGA 665  
DB 1923 GATGGGAAAGCTTCCCGCAAGCAAAAGCGCCCTAAGAGCTATTCGARAATTGGGA 1982  
QY 666 MCAATTTGACCTTCAGACCTAAGAAAGAAAGCACTATATTTCTTGCAGTCCGCTG 725  
DB 1983 CCAATTTGACCTTCAGACCTAAGAAAGAAAGCACTATATTTCTTGCAGTCCGCTG 2042  
QY 726 GCACTCCTGAGGGAAGTAAATTAACAACACTTACAGCTAGACTCTTTTGTGAA 785  
DB 2043 GCACTCCTGAGGGAAGTAAATTAACAACACTTACAGCTAGACTCTTTTGTGAA 2102  
QY 786 AAGGCAATAGAGTGAAGTCCATTAAGTAACTTCTTTCTTAAGAGCAACTCAC 845  
DB 2103 AAGGCAATAGAGTGAAGTCCATTAAGTAACTTCTTTCTTAAGAGCAACTCAC 2162  
QY 846 AATTATGTAATAAGTGAATTTATGCCCTACAGAGACCTTCAGAGTCTACCTCCATC 905  
DB 2163 AATTATGTAATAAGTGAATTTATGCCCTACAGAGACCTTCAGAGTCTACCTCCATC 2222  
QY 906 CGAGCATCCCGGCTCTCTCCCAATTAAGGACCCCTTCAACCCAAATGCTGCA 965  
DB 2223 CGAGCATCCCGGCTCTCTCCCAATTAAGGACCCCTTCAACCCAAATGCTGCA 2282

QY 966 AAGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTCCATATTTCCCAATTATGA 1025  
DB 2283 AAGAGATAGACAAAGGGTAAACAGTGAACCAAGAGTCCATATTTCCCAATTATGA 2342  
QY 1026 CCCCTCCCAAGCAGTGGGAGGAAAGAGATTCGGCCCAAGCAGAGTGCATGCTTTTYY 1085  
DB 2343 CCCCT-CCAGAGTGGGAGGAAAGATTCGGCCCAAGCAGAGTGCATGCTTTTTC 2401  
QY 1086 TCTCCAGACTTAAGCAATTAATAAAGAGCTTATAGTAATTCAGATTAACCTGATG 1145  
DB 2402 TCTCCAGACTTAAGCAATTAATAAAGAGCTTATAGTAATTCAGATTAACCTGATG 2461  
QY 1146 CTATATGATGTTTTCAGAGGTTAGACATTTCTTGATCTGACATGAGAGATATATA 1205  
DB 2462 CTATATGATGTTTTCAGAGGTTAGACATTTCTTGATCTGACATGAGAGATATATA-A 2520  
QY 1206 TGTCACTGCTTAATTCAGACCTAACCCCAATGAGAGAGTGCACCATTACTGACCT 1285  
DB 2521 TGTCACTGCTTAATTCAGACCTAACCCCAATGAGAGAGTGCACCATTACTGACCT 2580  
QY 1266 GAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGATGATGACAAGAGG 1325  
DB 2581 GAGGTTTGGCGATCTCTGGTATCTCAGTCAAGTCAATGATGATGACAAGAGG 2637  
QY 1326 AAGGAAATGATTTCCCAAGGCGAGCAGCAGTCCAGTCTASACCTCATTTGGGAC 1385  
DB 2638 AAGGAAATGATTTCCCAAGGCGAGCAGCAGTCCAGTCTASACCTCATTTGGGAC 2694  
QY 1386 ACGAAATCAGTAACTGAGAGATTTGCTGACACATTTGCTAACTTGTGCTASAA 1445  
DB 2695 ACGAAATCAGTAACTGAGAGATTTGCTGACACATTTGCTAACTTGTGCTASAA 2753  
QY 1446 GGACTAAGGAAATCTAGAAAGAAATCTAAGAAATCTAATGATGCTCCACATTAACA 1505  
DB 2754 GGACTAAGGAAATCTAGAAAGAAATCTAAGAAATCTAATGATGCTCCACATTAACA 2812  
QY 1506 GGGGAAAGGAAAGAAATCTTACTGCTTTCTGAGAGACTTAAGGAGGCAATTGAGAGC 1565  
DB 2813 GGGGAAAGGAAAGAAATCTTACTGCTTTCTGAGAGACTTAAGGAGGCAATTGAGAGC 2871  
QY 1566 GTGCTCTCTGTCACCTGACTGCTTTCTGAGAGGCACTAATCTTAAGGCTTAATGCA 1625  
DB 2872 GTGCTCTCTGTCACCTGACTGCTTTCTGAGAGGCACTAATCTTAAGGCTTAATGCA 2931  
QY 1626 CTCAGTCAAGTGCAGATTAAG-AAAAAATCTTAAGAGTGCCTAGAGCCCGAGCAAA 1684  
DB 2932 CTCAGTCAAGTGCAGATTAAG-AAAAAATCTTAAGAGTGCCTAGAGCCCGAGCAAA 2991  
QY 1685 ACTTGAAGAACTCTATTTGAACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA 1744  
DB 2992 ACTTGAAGAACTCTATTTGAACCTTGGCACTTGGCACTTGGCACTTGGCACTTGGCA 3051  
QY 1745 AGGCGGAAAGGAAAGGAGATTAATAAAGGCGACCGCTTATGATCAGACCTCTAGG 1804  
DB 3052 AGGCGGAAAGGAAAGGAGATTAATAAAGGCGACCGCTTATGATCAGACCTCTAGG 3111  
QY 1805 CAGGAGACTTTGAGAGCTCTGGGAAAGGAAAGGCGCAATTTGAATGCTTAATAGG 1864  
DB 3112 CAGGAGACTTTGAGAGCTCTGGGAAAGGAAAGGCGCAATTTGAATGCTTAATAGG 3171  
QY 1865 GCTTGTCTCAAGTGGCTTCAAGAGCACTTTAAAAAAGATTTGCTCAAGTGAAGTAA 1924  
DB 3172 GCTTGTCTCAAGTGGCTTCAAGAGCACTTTAAAAAAGATTTGCTCAAGTGAAGTAA 3231  
QY 1925 CGGCTCTCTGTCATGAGCCCTTATTTCAAGGAAATCACTGGAAGGCCATCTGCCCCAG 1984  
DB 3232 CGGCTCTCTGTCATGAGCCCTTATTTCAAGGAAATCACTGGAAGGCCATCTGCCCCAG 3291  
QY 1985 GGAAGAGGCTTTTGAAGTCAAGAGCACTAAACAGATATCCAGAGAGAGAGAGG 2044  
DB 3292 GGAAGAGGCTTTTGAAGTCAAGAGCACTAAACAGATATCCAGAGAGAGAGAGG 3351  
QY 2045 TGCCTGGGCAAGCGCATGCCATGCACTCAAGAGCCTTGCTATGCTTGAAC 2104

Db 3352 TGCTGGGGCAGAGCCATCCCATGCAATCACCCCTCACAGAGCCCTGGGTAATGCTTGAAC 3411  
Qy 2105 ATTGAGGGCCAGAGGTTGTCTCTGGACACTGGTGGCTTTCTTAAGTTTACTCTTC 2164  
Db 3412 ATTGAGGGCCAGG--AGGTTGTCTCTGGACACTGGTGGGCTTTCTTAAGTTTACTCTTC 3470  
Qy 2165 TGTCCCGACACTGTCTCTCCAGATCTGTCACTATTCTGAGGGGTCCTTAAGAGCGCA 2224  
Db 3471 TGTCCCGACACTGTCTCTCCAGATCTGTCACTA--TCTGAGGGGGTCTC--TAAAGCGGGCA 3528  
Qy 2225 GTCACTGATCTTTCCTCCAGCACTTAAGTTAAGTGGAGGCTTATCTCTTTCAC 2284  
Db 3529 GTCACTGATTC--TCTCCAGCACTTAAGTTATG--ACTGGGAGGCTTATCTTCTTTCAC 3586  
Qy 2285 ATGCTTTCTTAATATGCTTGAAGAGCCCACTACCTTGTAGGGAGACATTTGACAA 2344  
Db 3587 ATGCTTTCTTAATATGCTTGAAGAGCCCACTACCTTGTAGGGAGACATTTGACAA 3646  
Qy 2345 AAGCAGGGGCCATTATACACTGAAATAGAGAGAGAAACACCGCTTTGTGTCCTG 2404  
Db 3647 AAGCAGGGGCCATTATACACTGAAATAGAGAGAGAAACACCGCTTTGTGTCCTG 3705  
Qy 2405 CTGAGAGAGAAATTAATCCGGAAGTGGGGCAACAGAGCAATATGAGCGAGCCAA 2464  
Db 3706 CTGAGAGAGAAATTAATCCGGAAGTGGGGCAACAGAGCAATATGAGCGAGCG 3764  
Qy 2465 GAATGCCGCTCTGTTCAAGTTAACTAAAGATTCACCTTCCCTTCCCTACAAAGCA 2524  
Db 3765 GAATGCCGCTCTGTTCAAGTTAACTAAAGATTCACCTTCCCTTCCCTACAAAGCA 3824  
Qy 2525 GTACCCCTCAGACCCAGGCCCAACAGATTCGAAAGATTTGTTAAGACTTAAAGC 2584  
Db 3825 GTACCCCTCAGACCCAGGCCCAACAGATTCGAAAGATTTGTTAAGACTTAAAGC 3884  
Qy 2585 CCAAGGCTTAATAACATGATTAATCCCTGCACTATCCGATCCGATGGAATGAGAG 2644  
Db 3885 CCAAGGCTTAATAACATGATTAATCCCTGCACTATCCGATCCGATGGAATGAGAG 3937  
Qy 2645 CACAGAAACCAAGTGAAGTGAAGGTTAGTCAAGATTCAGAGATTAATGAGAGC 2704  
Db 3938 TACAGAAACCAAGTGAAGTGAAGGTTAGTCAAGATTCAGAGATTAATGAGAGC 3995  
Qy 2705 CGTGTCTTTTATACCAAGCTGTAAGTCCCTTAATCTGTGATTTCCCAATACCA 2764  
Db 3996 TGTGTCTCTCTAATGCAAGCTGTACTAGCCCTTAATCTGTGATTTCCCAATACCA 4055  
Qy 2765 GGAAGCAAGTGTTAATASCTGCTGAAGTCTTGAAGATGCTTCTGATCCCTGTA 2824  
Db 4056 GGAAGCAAGTGTTAATASCTGCTGAAGTCTTGAAGATGCTTCTGATCCCTGTA 4115  
Qy 2825 TCCTGACTCTCAATCTTGTGCTTGAAGATCTTCAAAACCAACATCTCAATC 2884  
Db 4116 TCCTGACTCTCAATCTTGTGCTTGAAGATCTTCAAAACCAACATCTCAATC 4175  
Qy 2885 CTGACTTTTATACCCCAAGGTTCAAGGATGATGCCCATTAATTTGGCAGGCTTAC 2944  
Db 4176 CTGACTTTTATACCCCAAGGTTCAAGGATGATGCCCATTAATTTGGCAGGCTTAC 4235  
Qy 2945 CCAAGACTTGAATCAATCTGCTGAAGTCTTGTGCTTGAAGTGTGAGTGAATTT 3004  
Db 4236 CCAAGACTTGAATCAATCTGCTGAAGTCTTGTGCTTGAAGTGTGAGTGAATTT 4293  
Qy 3005 ACTTTGCGGCTTCAAGAACTTGAAGTCAAGCAGCAGCAGCTCTTAATTT 3064  
Db 4294 ACTTTGCGGCTTCAAGAACTTGAAGTCAAGCAGCAGCAGCTCTTAATTT 4353  
Qy 3065 CCTGCTGACTGTGCTACAGAGGTTCCAAAGSAPARCTGCTGCTGCTACAGAGGT 3124  
Db 4354 CCTGCTGACTGTGCTACAGAGGTTCCAAAGSAPARCTGCTGCTGCTACAGAGGT 4413  
Qy 3125 TAAATACTTAAGTCTAATTAATCCAAAGGCAAGAGGCTTCAAGGAGAAATATCA 3184

Db 4414 ----TACTTAGGGCTAAATATATCCAAAGGCAACAGGCGCCCTCAGTGAAGAACATCA 4469  
Qy 3185 GCTTAATATGCTTAATCTCAATCTCAAGAAACCCCTTAAGCACTAAGRATTTCTTGGCT 3244  
Db 4470 GCTTAATATGCTTAATCTCAATCTCAAGAAACCCCTTAAGCACTAAGGAGTCTTGGCT 4529  
Qy 3245 AATAGGTTCTGCGAATATGATTCCTCCAGTGTGCGAATATGAGGATTAATATA 3304  
Db 4530 AATAGGTTCTGCGAATATGATTCCTCCAGTGTGCGAATATGAGGATTAATATA 4588  
Qy 3305 CAAATATTAAGAACTCAGAAAGCAATACCAATTAATATGATGATGATGATGAT 3364  
Db 4589 CAAATATTAAGAACTCAGAAAGCAATACCAATTAATATGATGATGATGATGATGAT 4648  
Qy 3365 AATGCTTTCCAGAGCCCTTAAAGAGGCTTAAACCAAGCCCAATGTTAAATGCTC 3424  
Db 4649 AATGCTTTCCAGAGCCCTTAAAGAGGCTTAAACCAAGCCCAATGTTAAATGCTC 4694  
Qy 3425 AACAGGGCAGACCTTTTATATATATATATATATATATATATATATATATATATAT 3484  
Db 4695 AACAGGGCAGACCTTTTATATATATATATATATATATATATATATATATATATAT 4753  
Qy 3485 TTAACAGATCCAGAGGATGAGCTTGAACCTGTGCTGCTGCTGCTGCTGCTGCTGCT 3544  
Db 4754 TTAACAGATCCAGAGGATGAGCTTGAACCTGTGCTGCTGCTGCTGCTGCTGCTGCT 4813  
Qy 3545 TATGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3604  
Db 4814 TATGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4873  
Qy 3605 CTGAAGAGTTTAAAT 3664  
Db 4874 CTGAAGAGTTTAAAT 4933  
Qy 3665 AATGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3724  
Db 4934 AATGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4993  
Qy 3725 AGGCTCTTAATCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3784  
Db 4994 AGGCTCTTAATCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5053  
Qy 3785 YCNCAATTTCTGCAAGATGAGAGAAAGATTAATATATATATATATATATATATATAT 3844  
Db 5054 YCNCAATTTCTGCAAGATGAGAGAAAGATTAATATATATATATATATATATATATAT 5113  
Qy 3845 AATCTATGCTGCTGAGAGGCTTATGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3904  
Db 5114 AATCTATGCTGCTGAGAGGCTTATGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 5172  
Qy 3905 TTTGATATCTGATGAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3964  
Db 5173 TTTGATATCTGATGAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5232  
Qy 3965 GTCAATGATTAATGAGATTAATGAGATTAATGAGATTAATGAGATTAATGAGATTAAT 4024  
Db 5233 GTCAATGATTAATGAGATTAATGAGATTAATGAGATTAATGAGATTAATGAGATTAAT 5292  
Qy 4025 GCAGAACTAATAGCTTCAATGAGGCTTCAATGAGGCTTCAATGAGGCTTCAATGAGG 4084  
Db 5293 GCAGAACTAATAGCTTCAATGAGGCTTCAATGAGGCTTCAATGAGGCTTCAATGAGG 5352  
Qy 4085 TATACAGCTTAT 4144  
Db 5353 TATACAGCTTAT 5412  
Qy 4145 AGGAATTTCTTAATCTTGAAGGATCACTTAATCAATCACTTAATCAATCACTTAATCA 4204  
Db 5413 AGGAATTTCTTAATCTTGAAGGATCACTTAATCAATCACTTAATCAATCACTTAATCA 5472  
Qy 4205 TTAATGCTGATGAGAACTTAATGAGGCTTCAATGAGGCTTCAATGAGGCTTCAATGAG 4264  
Db 5473 TTAATGCTGATGAGAACTTAATGAGGCTTCAATGAGGCTTCAATGAGGCTTCAATGAG 5532



4265 GGAAGRAAAGGAAATASAAGRAAYTGCCAAGCAATATTTGAAGCAAAGAAGCTGCA 4324  
4325 AGGACAGACCCCTCATTTAGAAATGCTTATTTAACTTCCCTTATAGATAGGGTAAATCCCTTC 4384  
5593 AGGACAGACCCCTCATTTAGAAATGCTTATTTAACTTCCCTTATAGATAGGGTAAATCCCTTC 5652  
4385 CGGGAACCAAGCCCGCTACTGACGAGAGAAACAAATGAGGAACTTACGAGG- CAG 4443  
5653 CGGGAACCAAGCCCGCTACTGACGAGAGAAACAAATGAGGAACTTACGAGG- CAG 5712  
4444 TTTTCTCCCTCGGGAAGCTTGAAGCACTGAAGAAAGGAAATATCTTTGCTGCAATAT 4503  
5713 TTTTCTCCCTCGGGAAGCTTGAAGCACTGAAGAAAGGAAATATCTTTGCTGCAATAT 5772  
4504 CCAATGGAATTTACTTTAAACCTTTCATCAACCTTTCACTTGAAGCATGATGACACCA 4563  
5773 CCAATGGAATTTACTTTAAACCTTTCATCAACCTTTCACTTGAAGCATGATGACACCA 5832  
4564 TCARATGGCCAAATCATATTTATCTGACAGGCTTTCAAAAATACTCAAGABAAAT 4623  
5833 TCAGATGGCCAAATCATATTTATCTGACAGGCTTTCAAAAATACTCAAGABAAAT 5892  
4624 CAGGCTGTGAATGTGACABAAATAATCCCTGCTATGCGCAAGCTCCTTACAG 4683  
5893 CAGGCTGTGAATGTGACABAAATAATCCCTGCTATGCGCAAGCTCCTTACAG 5952  
4684 ARAACAARAAGAGGCTTATCCCTGABARAACCTGCAACTGATTTTACCACAGCC 4743  
5953 AGAACAARAAGAGGCTTATCCCTGABARAACCTGCAACTGATTTTACCACAGCC 6012  
4744 ARAACAARAAGAGGCTTATCCCTGABARAACCTGCAACTGATTTTACCACAGCC 4803  
6013 ARAACAARAAGAGGCTTATCCCTGABARAACCTGCAACTGATTTTACCACAGCC 6072  
4804 GCTTCCCTGTAGAGCAGAAAAAGCCCAAGAGTAATAAGGACATGTTTCAATGAATA 4863  
6073 GCTTCCCTGTAGAGCAGAAAAAGCCCAAGAGTAATAAGGACATGTTTCAATGAATA 6132  
4864 ATTCCAGATTTGGAATTTCCCGAGGCTTACAGATGACATAGCCCTGCTTCCAGCC 4923  
6133 ATTCCAGATTTGGAATTTCCCGAGGCTTACAGATGACATAGCCCTGCTTCCAGCC 6192  
4924 ACGATTAACCGAGAGATACCGAGGCTTATGATTAAGATACATTAACCTGAGGCTG 4983  
6193 ACGATTAACCGAGAGATACCGAGGCTTATGATTAAGATACATTAACCTGAGGCTG 6252  
4984 AGGCAACAGTCTCAGGGAAGTGAAGAAATGATGAATGAATGAATGAATGAATGAAT 5043  
6253 AGGCAACAGTCTCAGGGAAGTGAAGAAATGATGAATGAATGAATGAATGAATGAAT 6312  
5044 AAGCAAAACCGAGAAACCACTTCAATGAGCTGCTGCTGCTTATGCTTAAAAA 5103  
6313 AAGCAAAACCGAGAAACCACTTCAATGAGCTGCTGCTGCTTATGCTTAAAAA 6372  
5104 ATTGCAACTTTCCCAAAAAGAGAGATTAAGCCCAAGAAATGCTGATGAAGGCC 5163  
6373 ATTGCAACTTTCCCAAAAAGAGAGATTAAGCCCAAGAAATGCTGATGAAGGCC 6432  
5164 TTGATTAACCAATGCTTGTGCTTGAACCAAGACAGCACTTATGAGACATCACT 5223  
6433 TTGATTAACCAATGCTTGTGCTTGAACCAAGACAGCACTTATGAGACATCACT 6492  
5224 CTTAGCAAAATATCAAAAGTTCTTAAAAATTACAAGAACTTATCCTGAGAGAG 5283  
6493 CTTAGCAAAATATCAAAAGTTCTTAAAAATTACAAGAACTTATCCTGAGAGAG 6552  
5284 GAAAAAATCTATTCACCCMGTGACATGATTAAGTCAAGTCCCTTCCTTAAATCC 5343  
6553 GAAAAAATCTATTCACCCMGTGACATGATTAAGTCAAGTCCCTTCCTTAAATCC 6612

5344 CATCCCTAGATACATCTGGGAGAGACCTTACCAAGTCAATTTATATATCCCAATGCGG 5403  
6613 CATCCCTAGATACATCTGGGAGAGACCTTACCAAGTCAATTTATATATCCCAATGCGG 6672  
5404 TTAAATGCTGAGAGTGTGCTTGAATACATCACTTTGATCAAAATCTGATATGCTC 5463  
6673 TTAAATGCTGAGAGTGTGCTTGAATACATCACTTTGATCAAAATCTGATATGCTC 6732  
5464 CAAAGAACCTGAAAAATCCAGAGAACAGTACGATTTCTGTGAACCTCTAGAGATT 5523  
6733 CAAAGAACCTGAAAAATCCAGAGAACAGTACGATTTCTGTGAACCTCTAGAGATT 6792  
5524 TGGCTGCTCTTCAACCAACCAAGAGAAAGTAAATCAATAATCCCATG 5583  
6793 TGGCTGCTCTTCAACCAACCAAGAGAAAGTAAATCAATAATCCCATG 6851  
5584 GGCCTCCCTTATCATATTTTCTCTATGTTTATGCTTATGCTTATGCTTATGCTTAT 5643  
6852 GGCCTCCCTTATCATATTTTCTCTATGTTTATGCTTATGCTTATGCTTATGCTTAT 6911  
5644 CCCCCTCATGCGCTGTATGACCAAGTACCTCCCTACCAAGTCTTATGAGAAATG 5703  
6912 CCCCCTCATGCGCTGTATGACCAAGTACCTCCCTACCAAGTCTTATGAGAAATG 6971  
5704 CAGGCTCCGGAATATTTGATGCTCCCATGCTATGAGAGTCTTATGAGAAATG 5763  
6972 CAGGCTCCGGAATATTTGATGCTCCCATGCTATGAGAGTCTTATGAGAAATG 7031  
5764 TTCACTGCGCAACCAATATGCTGCGCACTGCTATCACTGCGCACTGCTGATGAT 5823  
7032 TTCACTGCGCAACCAATATGCTGCGCACTGCTATCACTGCGCACTGCTGATGAT 7091  
5824 GCAAAATCATATATGAGACAGAAAAATGATTAATCTATGTTGCTGAGAGACTTGA 5883  
7092 GCAAAATCATATATGAGACAGAAAAATGATTAATCTATGTTGCTGAGAGACTTGA 7151  
5884 GTCACTGCTGTTGACTTATCTTACCCAACTGCTATGCTGATGAGGCTGAGATTCA 5943  
7152 GTCACTGCTGTTGACTTATCTTACCCAACTGCTATGCTGATGAGGCTGAGATTCA 7211  
5944 GATCAGGCAAGAAAAATGATTAAGAAAGTATCCCACTGCGGCTGATGATGCTG 6003  
7212 GATCAGGCAAGAAAAATGATTAAGAAAGTATCCCACTGCGGCTGATGATGCTG 7271  
6004 ACTCTAGCCCTTACAAAGAGTATCTCTCAAACTATCATGAACCTCTGATGATG 6063  
7272 ACTCTAGCCCTTACAAAGAGTATCTCTCAAACTATCATGAACCTCTGATGATG 7331  
6064 ACTGCTGCTGTAAGCTTATTAATACACCTCTCACTGAGCTCATGAGGCTGAGGCT 6123  
7332 ACTGCTGCTGTAAGCTTATTAATACACCTCTCACTGAGGCTCATGAGGCTGAGGCT 7391  
6124 AACCTTACATGCTTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6183  
7392 AACCTTACATGCTTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7451  
6184 GTACTGAAACAAATGAAACAACTTACAGCAAGAAATGAACCACTTCCGTTTATGAG 6243  
7452 GTACTGAAACAAATGAAACAACTTACAGCAAGAAATGAACCACTTCCGTTTATGAG 7511  
6244 CTTCTGTTTCCAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 6303  
7512 CTTCTGTTTCCAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 7571  
6304 AATTAACAATGAAACAACTTACAGCAAGAAATGAACCACTTCCGTTTATGAG 6363  
7572 AATTAACAATGAAACAACTTACAGCAAGAAATGAACCACTTCCGTTTATGAG 7631  
6364 GTCTGCTACCTCAGAAATATTTTGTCTGTGATGCTGATGCTGATGCTGATGCTG 6423  
7632 GTCTGCTACCTCAGAAATATTTTGTCTGTGATGCTGATGCTGATGCTGATGCTG 7691  
6424 GGTCTTCAAAATCTATGCTTCTCTCATTTCTATGAGGCTCCCTATGACATCTACCT 6483

Db	7692	GGCTCTTCAGAAATCTATGTCCTTCCCTCAATTTTAGTGCCCTTATGACCATCTACACT	7751
Oy	6484	GAAACAAGATTATATACAGTTATGTCTAATATCTAAGCCCCCGCAAAAGAAGTACCAATTCCT	6543
Db	7752	GAAACAAGATTATATACAGTTATGTCTAATATCTAAGCCCCCGCAAAAGAAGTACCAATTCCT	7811
Oy	6544	CCCTTTTGTATATAGGACAGAGAGGCTAAGGGCACTAGGTACTGGCACTTGGCCGTATCA	6603
Db	7812	CCCTTTTGTATATAGGACAGAGAGGCTAAGGGCACTAGGTACTGGCACTTGGCCGTATCA	7871
Oy	6604	ACCTCTACTCAGTTCTACTACAAACATCTCAAGAACTAAATGGGGACATGAAACGGGTC	6663
Db	7872	ACCTCTACTCAGTTCTACTACAAACATCTCAAGAACTAAATGGGGACATGAAACGGGTC	7931
Oy	6664	GCCGACTCCCTGGTCCACTTGCAGAGTCAACTTAATCCTCCTAGACAGTATGTCCTCPA	6723
Db	7932	GCCGACTCCCTGGTCCACTTGCAGAGTCAACTTAATCCTCCTAGACAGTATGTCCTCPA	7991
Oy	6724	AATCGAAGGCTTTAGACTCTGCTCAACCGCTGAPAAAGGGGGAACTGTTATTTTAAAGG	6783
Db	7992	AATCGAAGGCTTTAGACTCTGCTCAACCGCTGAPAAAGGGGGAACTGTTATTTTAAAGG	8051
Oy	6784	GAAAGATGCTGTATTAATGTTATCAATCCGGAATGTGCTAGAGAAAGTTAAAGAAAT	6843
Db	8052	GAAAGATGCTGTATTAATGTTATCAATCCGGAATGTGCTAGAGAAAGTTAAAGAAAT	8111
Oy	6844	CSAGATCGAATCAACGTTAKAGCAGARAGACTTCGAAACATCTGACCCCTGCTCCTC	6903
Db	8112	CSAGATCGAATCAACGTTAKAGCAGARAGACTTCGAAACATCTGACCCCTGCTCCTC	8171
Oy	6904	AGCCATGAGATGCTCTGATATTCCTCCCTCTTATGACCTCTAGACGTATATATTTGCTA	6963
Db	8172	AGCCATGAGATGCTCTGATATTCCTCCCTCTTATGACCTCTAGACGTATATATTTGCTA	8231
Oy	6964	CTCCCTTTTGAACCCCTGTAATCTTBACTCTCTGTTAACTTTCTCTTCAAGATCGAA	7023
Db	8232	CTCCCTTTTGAACCCCTGTAATCTTBACTCTCTGTTAACTTTCTCTTCAAGATCGAA	8291
Oy	7024	GCTGTAAACCTACAAATGAGGCCAAGATGACATCGAAGCTAAGATCTACCGCAGACC	7083
Db	8292	GCTGTAAACCTACAAATGAGGCCAAGATGACATCGAAGCTAAGATCTACCGCAGACC	8351
Oy	7084	CTGACCCGGCTGTATGAGCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCTGAG	7143
Db	8352	CTGACCCGGCTGTATGAGCCACGATCTGATGTTAATGACATCAAAGGCACCCCTCTGAG	8411
Oy	7144	GAAATCTCAGCTGACAACTCTCTAATAGGCCCCCAATTCAGCAGAGAAAGCATTTAGACGGT	7203
Db	8412	GAAATCTCAGCTGACAACTCTCTAATAGGCCCCCAATTCAGCAGAGAAAGCATTTAGACGGT	8471
Oy	7204	SGTCGGCCAAACCTCCCAACAGACGACTTAGGTTTCTGTTGAGATGGGGGAC	7255
Db	8472	SGTCGGCCAAACCTCCCAACAGACGACTTAGGTTTCTGTTGAGATGGGGGAC	8523
RESULT 3			
/ Sequence 17417, Application US/09949016			
/ Patent No. 6812339			
/ GENERAL INFORMATION:			
/ APPLICANT: VENTER, J. Craig et al.			
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
/ FILE REFERENCE: CL001307			
/ CURRENT APPLICATION NUMBER: US/09/949, 016			
/ CURRENT FILING DATE: 2000-04-14			
/ PRIOR APPLICATION NUMBER: 60/241, 755			
/ PRIOR FILING DATE: 2000-10-20			
/ PRIOR APPLICATION NUMBER: 60/237, 768			
/ PRIOR FILING DATE: 2000-10-03			
/ PRIOR APPLICATION NUMBER: 60/231, 498			
/ PRIOR FILING DATE: 2000-09-08			

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17417
; LENGTH: 77772
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(77772)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17417

Query Match          39.5%; Score 2994.2; DB 4; Length 77772;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 3418; Conservative 149; Mismatches 362; Indels 69; Gaps 23;

QY      1812  ACTTTGAGAGGCTGTGAAAAAGGAAAAAGCTGGGCAATTGTAATGCTTAATAGGGCTTGCT 1871
DB      19931  ACTTTGAGAGGCTGTGAAAAAGGAAAAAGCTGGGCAAAATCAAAATGCTGTAGGGCTTGCT 19872

QY      1872   TCCAGTACGGGTCTACAAAGACACTTTAAAAAAGATTGTCCAGTGAATAAGCCGCCCC 1931
DB      19871   TCCAGTGTGTCTACAAAGACACTTTAAAAAAGATTGTCTGAATAGAAATTAAGCTGCCCC 19812

QY      1932   TTGGTCACATGCCCTTAATTTCAAGGAACTCATCTGAAAGGCCCACTGCCCCAGGGGCAAA 1991
DB      19811   -TATTCATATCCCCCTCATGTCTGAAGGAAATCATCTGAAAGGCCCACTGCCCCAGGATGAA 19755

QY      1992   GGTCTTTTGAATGACGAAGACCACTAACAGATGATCCAGACAGACAGACTAGAGGGTGTCTGG 2051
DB      19752   GGTCTTTTGAATGACGAAGACCACTAACAGATGATCCAGACAGACAGACTAGAGGGTGTCTGG 19693

QY      2052   GGCAGAGCGCATATCCATATGCCATACCCCTCACAGAGCCCTGGGGTATGCTTGAACATTGAGG 2111
DB      19692   GGCAGAGCGCAGCCCAT-----GCCCAGGATGCTTGACCATTTGGGG 19655

QY      2112   GCCAGGAAGGT---TGCTCTGTGACACTGTGTCGGGTCTTCTTAAGCTTAATCTTCTGTCTGTC 2168
DB      19649   GCCAGGAAGGTAACTGTCTCTGTGACACTGTGTCGGGTCTTCTTAAGCTTAATCTTCTGTCTGTC 19599

QY      2169   CCGGACAACTGTCTCCAGATCTGTCACTAATTTCTGAGGGGGTCCNTAAGACGGGCACTCA 2228
DB      19589   CTGGACAACTGTCTCCAGATCTGTCACTAATTTCTGAGGGGGTCCNTAAGATAGGCAGTCA 19533

QY      2229   CTAGATATCTTTTTCCTCCAGCCCACTAAGTATGAACTGGGGAGCTTTATCTTTTCAATGCTC 2288
DB      19532   CTAGATATCTTTTTCCTCCAGCCCACTAAGTATGAG-ACCTGGGGAACTTCAATCTTTTCACTTCC 19475

QY      2289   TTTTCTAATTAATCTTGAAGAGCCCACTACCTTGTAGGGAGAGACATTTAGCAAAAGC 2348
DB      19474   CTTTCTAATTAATCTTGAAGAGCCCACTACCTTGTAGGGAGATCAATTTAGCAAAAGC 19411

QY      2349   AGGGGCACTATATACACTGACATAGAGAGAGAACCCGTTTGTGTGTTGCCCTGCTG 2408
DB      19414   AGGGGCACTATATACACTGACATAGGGAGAGAGAACCCGTTTGTGTGTTGCCCTGCTG 19355

QY      2409   AGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGACATATGAGACGACAAAGAAAT 2468
DB      19356   AGAAGGAATTAATCTCTGAAGTCTGGGCAACAGAGACATATGAGATGAG-AGAAAGAAAT 19299

QY      2469   GCCCGTCTCTTCAAGTAAACTTAAAGAAATTCACATTTCCCTTCCCAACAAAGGCAATAC 2528
DB      19297   GCCCGTCTCTTCAAGTAAACTTAAAGAAATTCGCGCTCTCTTCCCAACAAAGGCAATAC 19233

QY      2529   CCCCCTGAGACCAAGGCCCAACAAGGATTCCAAAGATTTGTTAAGACTTAAAGCCCAA 2588
DB      19237   CTTTCTGAGACCAAGGCCCAACAAGGATTCCAAAGATTTGTTAAGACTTAAAGCCCAA 19177

QY      2589   GGCTTAAATAAACATGCAATACCTCCCTGCAATATTCGGTATGGAATTGAGAGGCAACA 2648
DB      19177   GGCTTAAATAAACATGCAATACCTCCCTGCAATATTC-----AATTTAAGAGATAC 19121

QY      2649   GAAACCAATGAGCACTGAGAGGTTATGTCAGAACTTCAGAGATTAATCAATGAGGCGCTT 2708

```



Oy	4865	TTCCAGATTGGACCTTCCCGAGGCTTAAAGATGACATTAACCTCGCTTCCAGGCGCA	4924
Db	16911	TTCCAGATTGGACCTTCCCGAGACTTACAGATGACATTAAGCTGCTTCCAGGCGCG	168532
Oy	4925	CAGTACCACGAGGATGCCAGCG---TTAGATATACGATATACATTACCTGCGCC	4980
Db	16851	CAGTATACCACGAGGATATCGCAGACGTTACTTATGAGGCAATACATATCACTTACCTGCGCC	167922
Oy	4881	TGAAGGCGCACAGTCCCTCAGGAAAGGTGAGAAAAATGAAATACATCAAGACATCTA	5040
Db	16791	TGAGGCGCACATCTCTCAGGAAAGGTGAGAAAAATGAAACATCAACATCTCA	167322
Oy	5041	AAAAAGCAAAACCGAGAAAACCACTCACAATGGCGCTGTCTGTGCGCTATAGGCTTAAAA	5100
Db	16731	AAAAAGCTACTCAGGAAACCCACTGCGATGCGCTGTCTGTGCGCATAGCTTATTA	166722
Oy	5101	AGAACTGTGCACTTTCCCAAAAAAGCAGACTTAAGCCATACGAAAAATGCTGTATGGAGG	5160
Db	16671	AGAAATACAAATCTCCCCCAAAAAAGCAGACTTAAGCCATACGAAATGCTGTATGGACG	166122
Oy	5161	CCCTTCATTAACCATGACCTTGTGCTTGAACCA-AGAAGACCACTTAATGTGAGACATC	5219
Db	16611	CCCTTCCTAACCAATGACCTTGTGCTTGAACCGAGAGCGCCAAATTAATGTGAGACATC	165522
Oy	5220	ACCTCCTTAGCCAAATATACAAAGATCTTAAAAACATTAACAAGAACTATCCCTGAGAA	5279
Db	16551	ACCTCCTTAGCCAAATATACAAAGATCTTAAAAACATTAACAAGAACTGTATCCGAGAG	164922
Oy	5280	GAGGAAAAAGAACTATTTCACACCCMWGTGACATGATTAATTAAGTCCCTTCYTAAAT	5339
Db	16491	GAGGAAAAAGAA---ATTCCACCCGTGATACATGATTAATTAAGTCCCTTCYTAAG	164352
Oy	5340	TCCCCATTCCTAATACATCTCTGGAAAGACCTTAACCAATCATTTTATTTACCCCACT	5399
Db	16434	TCCCCACCCCTAGATATACATCTCTGGAAAGACCCCTAACCAATCATTTTATCTAACCTGAAT	163752
Oy	5400	GGCGTTAAAGGGGTGGAGTGGAGATCTTGGATACATCACCTTGAGTCAAAATCCTGGATA	5459
Db	16374	GGAGTTAAAGTGTCTGGAGTGGAGTCTTGAACACA-----TCAAACCTTGATA	163262
Oy	5460	CTGCCAAAGAACTGTAAAAATCCAGAGACAAACGCTATGCTATTCCTGTGAACCTCTAGAG	5519
Db	16325	CCGCAAAAGAAACCCAAATAATTCAGAGACAAACGCTATGCTATTCCTGTGAACCTCTAGAG	162662
Oy	5520	GATTTGCGCGTGTCTTCAACAAACAACAGAGGAAAGTAACTAAATATATAATCCCC	5579
Db	16265	GATCTGACACGTGCTTCTCAACCGACAACCTGGAGGAAAGTAACTAAATATATGAGACTCC	162062
Oy	5580	CATGAGCCTCCCTTATCATATTTTCTCTATASGTSTTTA-CCCTSTTTACACTCA	5638
Db	16205	--ATGAGCCCCCGCTCATATTTTCTCTTATAGTGTCTTACCCCCCTTACACTATCA	161482
Oy	5639	CTGCACCCCTTCACATGCGCGTGTATGACACAGTACTCCCTTACCMAGAGTTCTATGGA	5698
Db	16147	CTGCACCCCTTCATGCGCGTATATCTACAGTACTCCCTTACCAAGAGCTCTATGGA	160828
Oy	5699	GAATGCACGTCCTCCGGAATATTTGATGCGCCATCTGTATAGAGTCTTTSATAGGAAACC	5758
Db	16087	GAATGCGGCTTCCCAAGAAATATGATGCGCCATGTATAGAGATTTTCTTAAAGGAAACC	160282
Oy	5759	CAACTTATCTGCCACACCCATATGCCCCGCAACTCGC	5796
Db	16027	CAACTTACACATCACACCCATATGCCCCCTGCACCTTC	15990

RESULT 4  
US-09-949-016-12249/C  
; Sequence 12249, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12249
; LENGTH: 77997
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ..(77997)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12249

Query Match          39.5%; Score 2994.2; DB 4; Length 77997;
Best Local Similarity 85.5%; Pred. No. 0;
Matches 3418; Conservative 149; Mismatches 362; Indels 69; Gaps 23;

```

QY	1812	ACCTTGGAAGGCTCTGGAAAAGGGAAGCTGGGCAAAATGTAATGCTATAGGCGTTGCT	1871
Db	19935	ACCTTGGAAGGCTCTGGAAAAGGGAAGCGTGGGCAATCAATGCTGCTATAGGCGTTGCT	19876
QY	1872	TCCAGTGCAGGTCTACAGGACACTTTAAAAAGATTGTCCAGTGAAGTAAAGCGGCCCC	1931
Db	19875	TCCAGTGTGTCTACAGGACACTTTAAAAAGATTGTCTGAATGAAATTAAGCTGGCCCC	19816
QY	1932	TTTCGTCAATGCCCCCTTAATTTCAAGGGAATCACTGGAAAGGCCCACTGCCAGGGGACAA	1991
Db	19915	-TATTCATTCGCCCTCATGTCTCAAGGGAATCACTGGAAAGGCCCACTGCCAGGGGATGA	19757
QY	1992	GGTCTTTTGAATCAAGAACCTAACCAATGATCTCAGCAGCAGGACTGAAGGTGCTTGG	2051
Db	19756	GGTCTTTTGAATCAGTGAAGGCACTAACCAATTAAGCAGCAGGAGGAGGAGGTGCCGG	19697
QY	2052	GGCAAGGCCATCCCATGCGATCACCCCTCAGAGAGCCCTGGGATGTTGTTGACATTGAG	2111
Db	19696	GGCAAGGCCACCGCCAT-----GCCCAAGTAATGCTTGACATTTGGGG	19654
QY	2112	GCCAGGAAGT---TGTCCTCTGGAACACTGTGCGGTCTTCTTAATCTTACTCTTCTGC	2168
Db	19653	GCCAGGAAGTTAACTGTCTCTCTGGAACACTGGCATGGCTTCTCAATCTTAATCTCTGCTC	19594
QY	2169	CCGGAACAATGTCTCTCCAGATCTGTCCACTATTTCTAGAGGGGTCMTAAAGACGGGAGTCA	2228
Db	19593	CTGGACAAACGTCCTTCCAAATCTGTCACTAACCC---GAGGGGTCTTAGAATAGGCAATCA	19537
QY	2229	CTAGATCTTTTTCAGAGCACTAAGTATGAACTGGGAGGCTTATCTTTTTCACATGC	2288
Db	19536	CTAGATTC-TTCTCCAGCACTAAGTGTG-ACGGGGAACTTCATCTTTTCACTTGC	19479
QY	2289	TTTTCTAATATATGCTTGAAGCCCACTACTCTTGTTAAGGAGAGACTTCTAGCAAAAGC	2348
Db	19478	CTTTCTAATATATATGCTTGAAGCCCACTACTCTTGTTAAGGAGATCAATCTAGCAAAAGC	19419
QY	2349	AGGGGCATATATACCTGTAACATAGAGAAAGAAACACCCTTGTATTGTTGCCCTGCTTG	2408
Db	19418	AGGGGCATATATACCTGTAACATAGGAAAGAAACACCGTTTGCT--GTCCCTACTTG	19361
QY	2409	AGGAAGAAATTAATCTCTGAAGTCTGGGCAACGAAGACAAATATATGAGAGCAAGAAAT	2468
Db	19360	AGGAAGAAATTAATCTCTGAAGTCTGGGCAACGAAGACAAATATGAGATAG-AGAAAGAT	19302
QY	2469	GCCCGTCTGTGAATTAATCTAAAGATTCACATCTTCTCCCTACCAAGCAGTAC	2528
Db	19301	GCCCGTCTGTGAATTAATCTAAAGATTCGCGTCTCTTCCCAACAAAGCAGTAC	19242

QY	25252	CCCCCTCAGGCCCAAGGCCCAACAGAGATTCCAAAAGATTGTTAAGCATTTAAAGCCCA	2588
Db	19241	CCTCTTAGACCCGAG3CCCAACAAAGACTCCAAAAGATTGTTAAGCATTTAAAGCCCA	19182
QY	2589	GGCTTAGTAAACCATGCAATACTCCCTGCAGTAATTCGTAAGTGAATTAGAGAGGACAC	2648
Db	19181	GGCCCTAGTAAACCATGCAATAAGCCCTGCAGTAATTCGTAAGTGAATTAGAGAGGACAC	19122
QY	2649	GAACCCCACTGGAACAGTGAAGGTTAGTGCAGAACTTCAGAGATTATCAATGAGAGCCGTT	2708
Db	19128	GAACCCCACTGGAACAGTGAAGGTTAGTGCAGAACTTCAGAGATTATCAATGAGAGCCGTT	19071
QY	2709	GGCCCTTAGTAAACCCAGCTGTACCTAGCCCT---TATCTGTAAGTTCCCAATACAGAG	2765
Db	19070	GTTCTCTTAGTAAACCCAGCTGTACCTAGCTGTACCTAGCTGTACCTGTCCAAATACAGAG	19011
QY	2766	GAAGCAGAGTGGTTTACASTCCTGGACCTTAGAGATCCCTCTTGCACTCCCTGTACAT	2825
Db	19010	GAAGCAGAGTGGTTTACAGTCTGTGACCTTAGAGATCCCTCTTGCACTCCCTGTACAT	18951
QY	2826	CCTGACTCTCAATTCTTGTGCTTGAAGATCTTCAAAACCACATCTCACTCAC	2885
Db	18950	CCTGACTCTCAATTCTTGTGCTTGAAGATCTTGAATCTTAAGTCTCAACACAC	18891
QY	2886	TGCACTRTTTTACCACCAAGGTTACAGGATATGTCGCCCATCTAATTGGCCAGCATTAAGCC	2945
Db	18890	TGCACTRTTTTACCACCAAGGTTACAGGATATGTCGCCCATCTAATTGGCCAGCATTAAGCC	18831
QY	2946	CAGACTTGAGYCAATWTTCATACCTGGACACTCTTGCTCTTGRTAKGTGATGATTA	3005
Db	18830	CGAGCTTGAGCCACAGTCTCTCATCTTGAGGCACTCTTGCTCTTGRTAGTGAATGATTA	18771
QY	3006	CTTTTGGGCGCYRTGCAAAACCTGTGGCATCAACCAACCAAGGRCCTTMAATTC	3065
Db	18770	CTTTTGGGCAACCATTCAGAACTTTGTGCATATAGCCACCAAGTGCCTTAACTTC	18711
QY	3066	CTCGYACTGTGGCTACAWGTTTCCAAACSAARAGCTCARCTTGCTCAACAGAGTT	3125
Db	18710	ATGGCAACTGTGGCTACAAAGTTTCCAGACCAAGGCTCAGCTGTGCTCAACAGAGTT	18651
QY	3126	AAATACCTTAGRCCTAARVTTTCCAAAGGACCAAGGCCCTCAGTAGAGAAATATCCAG	3185
Db	18650	AAATACCTTAGGCGCTAATATATCCAGACCAACAGGCCCTCAGTAGAGAAATATCCAG	18591
QY	3186	CCTATACTGCGTTATTCCTCATCYCAAAACCTTAAGCACTAAGRGRATCTTGSCRTA	3245
Db	18590	CCTATACTGCGTTATTCCTTATCCCAAAACCATTAAGCACTTAAGAGGTTCTTGSCATA	18531
QY	3246	AYAGGYTTCGCCGAAMATGATTCGCCAGGTTGGCRAATAGCCAGYCATTAATATC	3305
Db	18530	ACAGGCTTCTGCCAAMATGATTCGCCAGGTTGGCRAATAGCCAGYCATTAATATC	18472
QY	3306	ASTAATTTAGAAACTCAGAAAGCCATATCCATTTATTAAGATGGGYMCTGAAGYMA	3365
Db	18471	ACTAATTTAGAAACTCAGAAAGGATATCCATTTATTAAGAGACACTGTACAGACA	18412
QY	3366	AGTGGCTTCCAGGCCCTTAAAGAGGCTTTAAACCCAGYCCCAAGTGTAAAGTTGCCA	3425
Db	18411	AGTGGCTTCCAGGCCCTTAAAGAGGCGCTT---AATCCAAACCTTATGTTAAGCTTGCA	18354
QY	3426	ACRGGGCAAGACTTTSTTATATRTCAAGAAAAAACAGRAAYAGCTCTGGGAGCTCT	3485
Db	18353	AAGGGGCAAGACTTTCTTTATGTGTACAG-AAAAAACAGRAAYAGCTCTTGAGATCTCT	18295
QY	3486	TACACAGRTCCRAAGAGYAGGCTTGCAACCYRTGGCRYACTTAATTAAGAAAYATGATG	3545
Db	18294	TACACAGRTCCRAAGAGYAGGCTTGCAACCAATGAGCATCTGATGTAAAGAAACATGATG	18235
QY	3546	AGTGGCAAAAGGCTTGRTCTATTGTTTAAAGGGTATGTGTGCGATAGCACTYKTAGTATC	3605
Db	18234	AGTGGCAAAAGGCTTGRTCTATTGTTTAAAGGGTATGTGTGCGAGTGTGCACTGTAGTATC	18175

Qy	3606	TGAACAGTTAAATAAATAACAGGGGAGAGACTTACTAGTGTGGACATCTCAATGAKTGAA	3665
Db	18174	TGAACAGATTAAATAAGATACAGGGAAGAGACTTACTAGTGTGGATGCTCTCATGTATGGGA	18111
Qy	3666	YKCAATTCACACTGTCTAAAGAGACTTGTGGCTGTCAACAACAGTTTACTTAATATCA	3725
Db	18114	CGGCAATCTCACTGTCTAAAGAGACTTACGGCTGTCAACAACAGTTTGTCTTAATATCA	1805
Qy	3726	GGCTCTATTACTTGAARGGCCAGTGTGTCACATGTGCAACTGTGCAACTCTTAAACGAGY	3785
Db	18054	GGCTCTATTACTTGAARGGCCAGTGTGTCACACTGGCACTGTGCAACTCTTAAATCCAGC	1799
Qy	3786	CNCATTTTCTTCCAGACATGAAAGAAAGATBAAATATACGTGTCAACAAATTAATTTCTCA	3845
Db	17994	CACATTTTCTTCCAGACATGAAAGAAAGATBAAACATACGTGTCAACAGGTGATTTCTCA	1793
Qy	3846	AACCTATGCCACTGAGGGGACCTTATAGAGTGCCTTGAACCTGATCCYAGCCCTCACT	3905
Db	17934	AACCTATGCCCTGAGGGGACCTTATAGAGTTCCTTGAACCTGATCCCAACCTCACT	1787
Qy	3906	TGTATACTGATGGAAGTTCCTTTGTAGAAAAAGGACTTGCAAAAGYGGGTATGACATGG	3965
Db	17875	TGTATACTGAGGAAGTTCCTTTGTAGAAAAAGCACTTGCAAAAGCAGGGTATGACATGG	1781
Qy	3966	TCAGTGATATATGAAATATTTGAAGTAATCCCTCACTCCAGGAATATGTCTTGAAGTTTG	4025
Db	17815	TCAGTGAATTAAGAAATCTTGAAAGTAATCCCTCACTCCAGGAATATGTGCTCACTG	1775
Qy	4026	CAGAACTPAATAGCCCTCAATKGGGCACTAGAAATTAGAGAAAGAAAGGTYAAATATAT	4085
Db	17755	CAGAACTPAATAGCCCTCACTCAGGCACTPAATTTAGAGAAAGAAAGGTTAATATAT	1769
Qy	4086	ATACAGACTCTPATATGTCTYACCTAGTCNTCCATGCGCATGMRGCAATATGSARAGAA	4145
Db	17695	AACAGACTCTPAAGATGTCTTACCTAGTCCTCCATGCGCATGCGCAATATGAGAGAA	1763
Qy	4146	GGGAATTCCTTAACCTTCTGAGRGAACACTCTACAACTACAGAAAGCATTAAGBARATAT	4205
Db	17635	GGGAATTCCTTAACCTTCCGAGGGAACCACTTACAAATTCAGAAAGCCATTAGGAGATTAAC	1757
Qy	4206	TATYGGGCTGACAGAAACCTABAGAGGTGMAAGCTTACACCTGCGGGGTCATCANAAAG	4265
Db	17575	AATYGGCTGTACAGAAACCTPAABAGGTGGGAGCTTACACTGCGGGGTCATCAGAAAG	1751
Qy	4266	GAAAGRAAAGGGAATATSAAGRGAAATYCCAGACAKATATTTGAAGCMAAAAGAGCTGCA	4325
Db	17515	GAAAGRAAAGGGAATATGAAGGGAACCGGCAACAGATATCTGAAGCCMAAAAGCCGCA	1745
Qy	4326	GGCAGAGCCTCCATTGAAATGCTTATTTAACTTCCCTTATGATAGGTAATCCCTTCC	4385
Db	17455	GGTGGGACCCCTCATTTGAAATGCTTATTAAGAGACCCCTATGATAGGGTAAATCCCTGT	1739
Qy	4386	GGGAAACCAAGCCCACTACAGCAGGAGAAACAGAAATGGGGAACCTCACGAGGACATTT	4445
Db	17395	AGGAAACCAAGCCCACTACTCAAAAGAAATAGAGTGGGAACCTCACGAGGACATAGTT	1733
Qy	4446	TTCTTCCCTCGGAGCGGTTAGCACTGAAGAAAGGAAATACTTTGCTGTCAACTATCC	4505
Db	17335	TTCTTCCCTCAGATGTGCTAGCCACGCAAGAAAGAAATACTTTGCTGTGAGCTATACC	1727
Qy	4506	AATGGAATTACTTAAAAACCTTTCATCAAACTTTTCACTTAGGCATGATAGACCCATC	4565
Db	17275	AATGGAATTACTTAAAAACCTTTCATCAGACTTTTCACTTAGGCATGATAGACCTATC	1721
Qy	4566	ABATGGCAATATCTTATTTTACGTGACCAAGCCTTTTCAAAACTATCAAGCAATATKTC	4625
Db	17215	AGAAGGCAATATCTTATTTTACGTGATCAGAGCTTTTCAAACTATCAAGCAATATGTC	1715
Qy	4626	GGGCTCTGAAKTGTGTCARABAAATATATCCCTGTCTYATCCCAAGCTCCTTTCAGAR	4685
Db	17155	GGGCTTATGAATGTGTGCAAGAAATATATATCCCTGTCTTATGTGCCAAGCTCCTTTCAGAG	1709
Qy	4686	AACAAABAACAGGCCATTAACCTGTARABAAATCTGGCAACT-GATTTTACCACAAAGCCA	4744





OY	5258	ACAAGAAACCTTATGAGGAAGAAGGAAAAAATATTCCACCCMMGTGCATGTTT	5317
Db	606	AATAAGAACCTATCCTCGAAGAGAGAAAAGAACTATTTCACCTTGTGCATGGTAT	665
OY	5318	AGTCAAGTCCCTTCCTCTTAATTTCCCACATCCCTAGATACATCTTGGAAGACCCTACCC	5377
Db	666	AGTCAAGTCCCTTCCTCTTAATTTCCCACATCCCTBATHATCTCTGGAGAACCTTACCC	725
OY	5378	AGTCATTTTATTTACCCCCACTGCCGTTTAAAGTGGCTGGAGTGGAGTCTTGGATATCA	5437
Db	726	AGTCATTTTATTTACCCCCACTGCCGTTTAAAGTGGCTGGAGTGGAGTCTTGGATATCA	785
OY	5438	CACCTGAGCAAAATCTCGGATATCTGCCCAAAGAAACCTGAATAATCCAGAGAACCGCTAG	5497
Db	766	CACCTGAGCAAAATCTCGGATATCTGCCCAAAGAAACCTGAATAATCCAGAGAACCGCTAG	845
OY	5498	CTATTCTGTGAACCTCTAGAGGATTTTGCGCCTGCTCTTCAACCAACCAAGAGAGAA	5557
Db	846	CTATTCTGTGAACCTCTAGAGGATTTTGCGCCTGCTCTTCAACCAACCAAGAGAGAA	905
OY	5558	GTAACATAAATCATTAATCCCCCATAGSCTCCCTTATCATATTTTCTCTKTASTGTS	5617
Db	906	GTAACATAAATCATTAATCCCCCATAGSCTCCCTTATCATATTTTCTCTKTASTGTS	964
OY	5618	TTTTAACCTSTTTTCACTCTCACCTGACCCCTCCATGCGCTGTATGACCATAGTACCTCC	5677
Db	965	TTTTAACCTSTTTTCACTCTCACCTGACCCCTCCATGCGCTGTATGACCATAGTACCTCC	1024
OY	5678	CTVACOMAGATTTTATGGAAGATGACAGCGTCCCGGAATPATTTAGTSCCCACGATAT	5737
Db	1025	CTVACOMAGATTTTATGGAAGATGACAGCGTCCCGGAATPATTTAGTSCCCACGATAT	1084
OY	5738	GGAGTCTTTSTAAGGAAACCCCAACCTTCACTGCCCCACCCATATGCCCCGCACTGCT	5797
Db	1085	GGAGTCTTTSTAAGGAAACCCCAACCTTCACTGCCCCACCCATATGCCCCGCACTGCT	1144
OY	5798	ATCACTCTGCACTTTTTCATGATGATGCAATATCTCATTTATGACACGAGAAAAATGATTA	5857
Db	1145	ATCACTCTGCACTTTTTCATGATGATGCAATATCTCATTTATGACACGAGAAAAATGATTA	1204
OY	5858	ATCCAGTGTCTCGGAGGACCTGGAGTCACTGCTGTGTGACCTTACCTTCAACCAACTG	5917
Db	1205	ATCCAGTGTCTCGGAGGACCTGGAGTCACTGCTGTGTGACCTTACCTTCAACCAACTG	1264
OY	5918	GTATGTCTGATGGGGGTGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGATTA	5977
Db	1265	GTATGTCTGATGGGGGTGAGTTCAAGATCAGGCAAGAGAAAAACATGTAAAAAGATTA	1324
OY	5978	TCTCCCACTGACCSGGGTACATGGCACCTGTAGCCCTTACMAAGACATAGTCTTCAA	6037
Db	1325	TCTCCCACTGACCSGGGTACATGGCACCTGTAGCCCTTACMAAGACATAGTCTTCAA	1384
OY	6038	AACTTACATGAAGAACCTTCCTGATCCCATATCTGCGCTGTGTAAAGCTTATTTAATACCCCTCA	6097
Db	1385	AACTTACATGAAGAACCTTCCTGATCCCATATCTGCGCTGTGTAAAGCTTATTTAATACCCCTCA	1444
OY	6098	CTGGGCTCCATGAGGTCTCGGCCCAAAACCTTACATCTGTTGTGATNGCTCCCTCTGA	6157
Db	1445	CTGGGCTCCATGAGGTCTCGGCCCAAAACCTTACATCTGTTGTGATNGCTCCCTCTGA	1504
OY	6158	ACTTCARGCCATATGTTCATTCCTCTGTACTGAAACAATGGAAACACTTGACACAGAAA	6217
Db	1505	ACTTCARGCCATATGTTCATTCCTCTGTACTGAAACAATGGAAACACTTGACACAGAAA	1564
OY	6218	TAAACACACATTCCTTTTAGTAGGACCTCTTGTTCCTAATSTGGAAATAACCATACCT	6277
Db	1565	TAAACACACATTCCTTTTAGTAGGACCTCTTGTTCCTAATSTGGAAATAACCATACCT	1624
OY	6278	CAAAACCTCACCTGTGTAAATTTTAGCAATATCATATACACAACAACCTCCCAATGATCA	6337
Db	1625	CAAAACCTCACCTGTGTGTAAATTTTAGCAATATCATATACACAACAACCTCCCAATGATCA	1684
OY	6338	GGTGGGTAACTCTCCCAACAAATAGCTGCTTACCTCAGAAATATTTTTGTCTGTG	6397
Db	1685	GGTGGGTAACTCTCCCAACAAATAGCTGCTTACCTCAGAAATATTTTTGTCTGTG	1744
OY	6398	GTACCTCAGCCCTATCGTGTGTTTGAATGGCTCTTGAGATCTATGTGCTTCCCTCATCTCT	6457
Db	1745	GTACCTCAGCCCTATCGTGTGTTTGAATGGCTCTTGAGATCTATGTGCTTCCCTCATCTCT	1804
OY	6458	TAGTCCCCCTATATGRCATCTTACACTGAAACAAGATTTATACGTTATGTCAATATTAAGC	6517
Db	1805	TAGTCCCCCTATATGRCATCTTACACTGAAACAAGATTTATACGTTATGTCAATATTAAGC	1864
OY	6518	CCCGCAAAAAGATACCCATTCCTCCTTTTGTATATAGGACAGSGAGTGTAGGTGCAC	6577
Db	1865	CCCGCAAAAAGATACCCATTCCTCCTTTTGTATATAGGACAGSGAGTGTAGGTGCAC	1924
OY	6578	TAGTATCTGGATTTGGCCGTATCACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAG	6637
Db	1925	TAGTATCTGGATTTGGCCGTATCACAACCTCTACTCAGTTCTTACTACAAACTATCTCAAG	1984
OY	6638	AACTAATGGGGAACATGGAACGGGTGCGCACTCCGTGTGACCTTGCAMAAGTCACTTA	6697
Db	1985	AACTAATGGGGAACATGGAACGGGTGCGCACTCCGTGTGACCTTGCAMAAGTCACTTA	2044
OY	6698	ACTCCCTAGCAGCATGATCTTTTAAATGGAAGGCTTTTAAAGATCTGCTAACCCCTGARA	6757
Db	2045	ACTCCCTAGCAGCATGATCTTTTAAATGGAAGGCTTTTAAAGATCTGCTAACCCCTGARA	2104
OY	6758	GAGGGGAACCTGTTTATTTTATGGGGAAGATCTGTATATATGTTAATCAATCCGAAA	6817
Db	2105	GAGGGGAACCTGTTTATTTTATGGGGAAGATCTGTATATATGTTAATCAATCCGAAA	2164
OY	6818	TGCTCACTGAGAAAATTAAAGAAATTSAGATGGAATACACGTAKACAGARGACTTC	6877
Db	2165	TGCTCACTGAGAAAATTAAAGAAATTSAGATGGAATACACGTAKACAGARGACTTC	2224
OY	6878	GAAACATGNAACCTGGGGGCTCCTCAGCCRAATGGAATGCGCTGGATTCTCCCTCTTAG	6937
Db	2225	GAAACATGNAACCTGGGGGCTCCTCAGCCRAATGGAATGCGCTGGATTCTCCCTCTTAG	2284
OY	6938	GACCTCTAGCAGCTAATAATATTGTCTACCTCCTTTGGAACCTGTATCTTTTACCTCTTG	6997
Db	2285	GACCTCTAGCAGCTAATAATATTGTCTACCTCCTTTGGAACCTGTATCTTTTACCTCTTG	2344
OY	6998	TTAATCTTGTCTTCCAGAAATCGAAGCTGTAAATCTAACAATGGAAGCCCAAGATGACGT	7057
Db	2345	TTAATCTTGTCTTCCAGAAATCGAAGCTGTAAATCTAACAATGGAAGCCCAAGATGACGT	2404
OY	7058	CCAAGACTBAATCTTACCGCAGACCCCTGGAACGGGCTGTATAGCCCAAGATCTGATGTTA	7117
Db	2405	CCAAGACTBAATCTTACCGCAGACCCCTGGAACGGGCTGTATAGCCCAAGATCTGATGTTA	2464
OY	7118	ATGACATGAAGGACCCCTCTGAGGAAATCTCAGCTGCAACCTCTACTAGGCCCA	7177
Db	2465	ATGACATGAAGGACCCCTCTGAGGAAATCTCAGCTGCAACCTCTACTAGGCCCA	2534
OY	7178	ATTGAGCAGGAAGATTTAGAGCGGTGTGCGCCCACTCCCAACAGCACTTAGTTTT	723

Accession	Sequence	Length
D5	AGAAATAGCCATCATCTATATGGCTGAGAGCAACAGACAGGAGGACAATATGCGGATATA	2824
Q7	AAACCAAGTCTTCGAGCGCGCAACGGCAACCCCTTTGGGTCCCTCCCTTTATAGGGA	7537
D5	AACCCAAAGCTTCGAGCGCGCAACGGCAACCCCTTTGGGTCCCTCCCTTTATAGGGA	2884
Q7	GCTCTGTTTCAFGCTATTCTCACTATTTAAATCTTGACCTGCR	7582
D5	GCTCTGTTTCACTGCTATTCTCACTATTTAAATCTTGACCTGCR	2929

```

RESULT 6
US-09-949-016-13002/c
/ Sequence 13002, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 13002
/ LENGTH: 168394
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(168394)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13002

```

	Query March 34.5%; Score 2615.4; DB 4; Length 168394;
Match Local Similarity 85.1%; Pred. No. 0;	
Batches 3333; Conservative 134; Mismatches 355; Indels 58; Gaps 33;	
QY	1012 TTCCCAATTAATGACCCCTCCCAAGCAGTGGGAGAGAAATTCGCGCCACGCAAGTG 1071
Db	29005 TCCCCCAAGGCAAAAACACCCCTAAATGATATTTGGAGAAATTCGCGCTGTCAAGTG 2894
QY	1072 CATGCGCTTTTYYTCTCCCAAGCTTAAGCAAAATAAAAACAGACTTGATGTAATTCGA 1131
Db	28945 TATGATACCTTTTCCCTGTCAACTTTGAAGCAAAATTTAAATAGACTTGGTAATTCGA 2888
QY	1132 GATAAATCTGATGCTATATTTGTTTACAAAGGTTAGACAAATCTTTGATCTGACA 1191
Db	28885 GATAACCCCTGTGGCTAATTTGATCTCTTACAAGGTTAGACAAATCTTTGATCTGACA 2882
QY	1192 TGGAGAGATATATATCTCACTGCTAAACAGACATAACCCCAATGAGAAAGTCCAC 1251
Db	28835 TGGAGAGATATA-ATGTTATCTGTAGATACGACACTAATCCCAATGAGAAAGTCTGC 2876
QY	1252 CATAACTGACGCTGAGAGTTTGGAGATCTCGTATCTCACTCAAGTCAATGATANGG 1311
Db	28766 CATAACTGACGCTGAGAGTTTGGAGATCTCGTATCTCACTCAAGTCAATG--ATAGG 28707
QY	1312 ATGACACAGAGAGAAAGAAATGATTTCCCAAGGCCACGACAGCAATTTCCACGCTASA 1371
Db	28708 ATGACACAG-AGGAAAAAGAACATTTCCCAAGGCCACGAGGAGATTTCCACAGTAGA 2865
QY	1372 CCGCTATGGGGAGACAGAAATCAGTACATGAGGAGATTGGTGTGACAGACTTTGCTAA 1431
Db	28659 CCGCTATGGG-----ACACAGATTCAGAAATGGAATTTGGTCCACAGACTTTGCTAA 2859
QY	1432 CTTGTGTCTASAGAGCTTAGAAAACTASGAAAGATCTTAYGAATTACTCAATGATG 1491

Db	28594	CTTGCTTGCTTA -AAGGACTTAAGGAAACTAAGAG -AGCCATTAAGATTATTCAGTAGT	28537
QY	1492	TCCACCAATAACAAGGGGAAAGGAAATAATCCTACTGCTTTCTGAGAGACTAAGGA	1551
Db	28536	TTCACTATAACA -AGAAAGGAAATAATCCTACTGCTTTCTGTAGACTAAGGGA	28478
QY	1552	GGCATTGAGGAAGCGTCTCTCTGACCTGACTCTTTGAAAGGCACTAATCTTAA	1611
Db	28477	GGCATTTGAAGGAAAC -AAACTCTCTGACCTGACTCTGTTGAAGGCCAATAATCTTAA	28415
QY	1612	GCGTAAGTTATCACTCAGTCAGCTSCAGACATTAG -AAAAAATCTCAAAAGTCGCGT	1670
Db	28418	GGTAAAGTTATATCACTCAGTCAGCTGCAACATTAGAAAAAATTTAAAGTCCACTT	28355
QY	1671	AGGCGCGAGCAAAACTTAAGAAACCTTATTTGAACCTTGCAACTGCGGTTTTTTATATAG	1730
Db	28358	AGGCGCGAGCAAAACTTAAGAAACCTTATTTGAACCTTGCAACTGCGGTTTTTTATATAG	28299
QY	1731	AGATCAGAGAGAGACAGGCGGAAACAGACAAAGGGATTTAAAAAAGCCACGCGTTAG	1790
Db	28298	AGATCAGAGAGAGACAGGCGGAAACAGATTAAGGGG -TAAAGAAAGGCCACACATTAG	28240
QY	1791	TCATAGACCCCTCAGGCAAG -TGAAGTTTGGAGGCTCTGAAAAAGGAAAAAGCTGGCAAT	1849
Db	28239	TCATGCGCCCTCAGGCAAGACACACTTCGGAAGCTCTGGAACGGGAAATCTGGACAAAT	28188
QY	1850	TGAATGCTTAATAGGGCTTGCTTCCAGTGGGCTTACAGAGACATTAAAAAAGTTGT	1909
Db	28179	TGAATGCTTAATAGGGCTTGCTTCCAGTGGGCTTACAGAGACATTAAAAAAGTTGT	28120
QY	1910	CCAAAGTGAAGTAAAGCC -GCCCTTGTCATAGCCCTTAATTTCAAGGGAAATCACTGGA	1968
Db	28119	TCTAATGAAATGAAGCCACCCCTGCTCATAGCCCTTAATTTCAAGGGAAATCACTGGA	28066
QY	1969	GGCCCACTGCCCGAGGGGACAAAGGCTCTTTGATCAGAGAGCCACTTAACAGATGATCCA	2028
Db	28059	GGCCCACTGCCCGAGGGGAGGAAGTCTCTGATCAGAGAGCCACTTAACAGATGATCCA	28000
QY	2029	GCAGCAGACTGAGAGGTGCTGCGGGCAAGCCATTCCTATGCTACCTTCACAGAGCC	2088
Db	27999	GCAGCAGACTGAGAGGTGCTGCGGGCAAGGCCAGCCACTGCAATCACTTCACAGAGCC	27944
QY	2089	CTGGGTATGCTTGAACCTTGAAGGGCCAGGAAGGT ---TGCTCTGAGCACTGCTGCGGT	2145
Db	27939	CCGGGTATGCTTGAACCTTGAAGGGCCAGAGGTTAATGTTCTCTTGAACACTGCGGTGCG	27888
QY	2146	CTTCTTAGTCTTACTCTCTGTGCCGGAACACTGCTCCAGTCTGTCACTAATTGAG	2205
Db	27819	CTTCTTAGTCTTACTCTCTGTCTGTCTTGAACAAGT -CTCAGATCTGTCACTATCCAAAG	2782
QY	2206	GGGGTCCNTAAGACGGGCACTCACTAGATACCTTATTCAGGCCACTAAGTTAAGAACTGG	2265
Db	27820	GG-----TCTGAGGACAGCCACTGATGATC -TTCCTCCAGCCACTAAGCTGTG -ACTGG	2777
QY	2266	GGACTTATTTCTTTTCAATGCTTTTCTAATATAGCTGAAGCCCACTACTCTGTA	2325
Db	27769	GGAACTTACTCTTTTCAATGCTTTTCTAATATAGCTTAACCCCACTCCCTGTTA	2771
QY	2326	GGGAGAGACATTTAGCAAAAGCAGGGGCCATTATACCTGCAACTAGGAGAGGACA	2385
Db	27709	GGGAAAGACATTTCAACAAAGCAGGGGCAATTATACCTGCAACTAGGAAAGGACA	27655
QY	2386	CCCGTTTGTGTNCCCTGCTTGAAGAAAGAAATTAATCTGAAGTCTGGGCAACAGAGG	2445
Db	27649	CCCGTTTGTGT -CCCTGCTTGAAGAAAGAAATTAATCTGAAGTCTGGGCAACAGAGG	2759
QY	2446	ACAATATAGACAGGCCAAAGAGCCGCTCTGTCAAGTTAACTAAAGATTCACCT	2505
Db	27590	ACAATATAGACAG -CAAAATTAATGCCGCTCTGTCTTAGTTAACTAAAGATTCACCT	2753
QY	2506	CCTTCCCTAACAAAGGCAATACCCCTCAGACCCAGGCCCAAGGATTCCAAAGA	2565

Db 27321 CTTTTCCCTACCAAGGAGTACCCCTTAGACCCGAGGCCCAACAGGATCCAAAAA 27472  
OY 2566 TTGTTAAGGACTTAAAGCCCAAGGCTTAGTAAACCATGCACTACCTCCCTCACTAATT 2625  
Db 27471 TTGTTAAGGACTTAAAGCCCTTCTCTAGTAAACCATGCACTAGCCCTGCAATCCCTC 27412  
OY 2626 CCGTAGTGATTTAGAGGACACAGAAACCAGTGCACAGTGGAGGTTAGTCAAGATCT 2685  
Db 27411 C-----AATTTTAGAGTAAAGAACCCATAGACAGTGA -GTTAGTCAAGAGT 27360  
OY 2686 CAGGATTTCATGAGAGGCCGTTCCTTTATACCAGCTGTACCTTAGCCTTACTG 2745  
Db 27359 CAGAGTTTCAGT -GAGCTGCTGTCTCTATACCAGCTGTACTTACCCTTACTC 27301  
OY 2746 TGMATTTCCCAATACAGAGGAGAGAGTGGTTTAACTCCCTGAGCCTTMAAGATGCT 2805  
Db 27300 TGCTTTCCCAATACAGAGGAGAGAGTGGTTTAACTCCCTGAGCCTTAAAGATGCT 27241  
OY 2806 TCTTCGATCCCTGTACATCCTGACCTCAATTCCTGTTGCTTGAAGATCTTCA 2865  
Db 27240 TTTTCTGATCCCTGTACATCCTGATTCCTCAATTCCTTATTTGCTTAAAGATGCTCA 27181  
OY 2866 ACCCAATCTCACTCACTGACCTTTTATCCCAAGGTTCAAGGATAGTCCCATC 2925  
Db 27180 ACCCAATGTCACACTCACTGACCTGTTTAAACCAAGGTTAAGATAGCCCATC 27121  
OY 2926 TATTTGGCAGGACTTGGCCCAAGCTTGAACATTTTCATTTTCACTGACACTTGTCC 2985  
Db 27120 TATTTGGCAGGACTTGGCTGAAGCTTGAACATTTTCATTTTCACTGACAGTGTGCC 27061  
OY 2986 TTCTGTAAGTGAATTAATTTTTCGTCGCTTTC -CAGAACTTGTGCACTCAAGC 3044  
Db 27060 TTTGTTAGTGAATTAATTTTCTAGTGTGCTTGAAGAACTTGTGCACTCAAGC 27001  
OY 3045 ACCCAAGCTCTTMAATTTCTGTCGTACTGTGCTTCAAGGTTTCCAAACATATCT 3104  
Db 27000 ACCCAAGCTCTTMAATTTCTGTCGTACTGTGCTTCAAGGTTTCCAAACATATCT 26941  
OY 3105 CACTCTGCTCAGAGGCTTAAATTAATTAAGGAACTCAAGAACTCAAGGAACTCAAG 3164  
Db 26940 CAGCTCTGCTTACAGAGGCTTAAATTAATTAAGGAACTCAAGAACTCAAGGAACTCAAG 26881  
OY 3165 CTCAGTGAAGAAATATCAAGCTTATCTGCTTATCTCATCTCACTCAAACTCAAG 3224  
Db 26880 CTCAGTGAAGAAATATCAAGCTTATCTGCTTATCTCATCTCACTCAAACTCAAG 26821  
OY 3225 CTAAAGRRRTTCTTGGCTTAATTAAGGAACTCAAGAACTCAAGGAACTCAAG 3284  
Db 26820 ATAAAG-GTTCCTTGGCATTAACAGGCTTTCGCAAAATATGATTT-CCAGGTACGCA 26763  
OY 3285 AATAGCCAGGCTAATTAATTAAGGAACTCAAGAACTCAAGGAACTCAAGGAACTCAAG 3344  
Db 26762 AATAGCCAGGCTAATTAATTAAGGAACTCAAGAACTCAAGGAACTCAAGGAACTCAAG 26703  
OY 3345 AAGATGAATTAAGGAACTCAAGGAACTCAAGGAACTCAAGGAACTCAAGGAACTCAAG 3404  
Db 26702 AAGATGAATTAAGGAACTCAAGGAACTCAAGGAACTCAAGGAACTCAAGGAACTCAAG 26644  
OY 3405 GYCCCACTGTTAAGTTGCCAAGGAGCAAGCTTTTSTTATTAATCAAGAAATTAAC 3464  
Db 26643 GACTCAGTGTAAAGCTTACCAAGGAGCAAGCTTTTCTTCAATATGTCAC--AAAAAC 26587  
OY 3465 AGAAATAGCTCTGAGGAACTCTTACAGGAACTCAAGGAACTCAAGGAACTCAAGGAACTCAAG 3524  
Db 26586 AGAAATAGCTCTGAGGAACTCTTACAGGAACTCAAGGAACTCAAGGAACTCAAGGAACTCAAG 26527  
OY 3525 CCGAATTAAGGAACTGATGATGAGGAACTGAGGAACTGAGGAACTGAGGAACTGAGGAACTGAG 3584  
Db 26526 CCGAATTAAGGAACTGATGATGAGGAACTGAGGAACTGAGGAACTGAGGAACTGAGGAACTGAG 26467  
OY 3585 GCGAGTGAAGTATTAATTAAGGAACTTAAATTAATCAAGGAACTGAGGAACTGAGGAACTGAG 3644  
Db 26466 GCGAGTGAAGTATTAATTAAGGAACTTAAATTAATCAAGGAACTGAGGAACTGAGGAACTGAG 26407

OY 3645 GTGACATCTTCATGAGGAAATGACATCTCACTGCTTAAAGGAACTTGTGCTGACGA 3704  
Db 26406 GTGACATCTTCATGAGGAAATGACATCTCACTGCTTAAAGGAACTTGTGCTGCTGACGA 26347  
OY 3705 CAACGTTTACTTAAATTCAGGCTTATTAATTAAGGAACTGAGGAACTGAGGAACTGAG 3764  
Db 26346 CAATCGTTTGTAAATTAATCAAGGCTTATTAATTAAGGAACTGAGGAACTGAGGAACTGAG 26287  
OY 3765 TTGTGCACTCTTAACCAAGGAACTTCTTCAAGCAATGAAGAAATGAATTAATTAATTAAT 3824  
Db 26286 TTGTGCACTCTTAACCAAGGAACTTCTTCAAGCAATGAAGAAATGAATTAATTAATTAAT 26229  
OY 3825 CTGTCAACATTAATTTCTCAAACTTACCACTGAGGAACTTGTAGAGTTCCTT 3884  
Db 26228 CTGTCAACATTAATTTCTCAAACTTACCACTGAGGAACTTGTAGAGTTCCTT 26170  
OY 3885 GACTGATCTGACCTTCAACTTGTATCTGATGGAATTCCTTGTAGAAATGAAGAACTTCA 3944  
Db 26169 GACTGATCTTAC-TCACTTGTATTAATGAATTCCTTGTAGAAATGAAGAACTTCA 26111  
OY 3945 GAAATGAGGATGACAGTGTGATGATTAATGAATTAATGAATTAATGAATTAATGAAT 4004  
Db 26110 -AAAGGAGGATGACAGTGTGATGATTAATGAATTAATGAATTAATGAATTAATGAAT 26052  
OY 4005 CAGGAATAGTGTGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 4064  
Db 26051 CAGGAATAGTGTGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 25992  
OY 4065 AAGAAATTAAGGAAATTAATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 4124  
Db 25991 AAGAAATTAAGGAAATTAATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 25932  
OY 4125 ATGTCATATGATGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 4184  
Db 25931 ATGTCATATGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 25872  
OY 4185 AAGAAATTAAGGAAATTAATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 4244  
Db 25871 AAGAAATTAAGGAAATTAATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 25812  
OY 4245 ACTGTCAGGATGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 4304  
Db 25811 ACTGTCAGGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 25752  
OY 4305 TTGAAGCAATGATGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 4364  
Db 25751 TTGAAGCAATGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 25692  
OY 4365 TAGTATAGGATTAATGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 4424  
Db 25691 TAGTATAGGATTAATGATGATGATTAATGAATTAATGAATTAATGAATTAATGAATTAATGAAT 25634  
OY 4425 GGGGAACTTAC--GAGGCAATTTCTCCCTCGGAGCGGTTAGCCATGAAGAAAGGA 4481  
Db 25633 GGGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAA 25574  
OY 4482 AATTAATCTTGTGCTGCACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAA 4541  
Db 25573 AATTAATCTTGTGCTGCACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAA 25514  
OY 4542 ACTTAAGCACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAA 4601  
Db 25513 ACTTAAGCACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAA 25454  
OY 4602 TCAAACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAA 4661  
Db 25453 TCAAACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAACTTACGAGGAA 25394  
OY 4662 CTATGCGCAAGCTCTTCA 4681  
Db 25393 ACTGCAAGGCAATTAATTTCA 25374

```
RESULT 7
US-09-949-016-15858/c
; Sequence 15858, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 15858
; LENGTH: 145320
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15858

Query Match      29.3%; Score 2220.4; DB 4; Length 145320;
Best Local Similarity 86.6%; Pred. No. 0;
Matches 2728; Conservative 92; Mismatches 216; Indels 115; Gaps 32;

QY      606 GATGGGAACGTTCCCGCCAGACAAAACGCCCTAGACGTATTCGGAAATTGGGA 665
DB      97137 GATGGGAACGTTCCCGCCAGACAAAACGCCCTAGACGTATTCGGAAATTGGGA 97078
QY      666 MCAATTTGACCCCTGAGACCTAAGAAAGAAAGCACTTATCTCTGCAAGTCCGCTG 725
DB      97077 CCAATTTGACCTCTGAGTGTCTAAGAAAAGAAAGAC--ATATCTCTCTGCAAGTCCGCTG 97020
QY      726 GCA-----CTCTGAGGAGATTAATTATTA 753
DB      97019 GCAACGATATACTCTTAAGGGGAGAAACCTGCGATCTGAGGAGATTAATTATTA 96960
QY      754 CACCATCTTAAGCTAGACCTTTTGTGAA--AAAGCAATGAGTGAATGCCCTA 810
DB      96959 CACCATCTTAAGCTAGACCTTTTGTGAA--AAAGCAATGAGTGAATGCCCTA 96900
QY      811 AGTCAAACTTCTTTCAATTAAGAGACAACTCACATTAATTAATAAGTGTATTA 870
DB      96899 GGTCAAACTTCTTTCAATTAAGAGACAACTCACATTAATTAATAAGTGTATTA 96840
QY      871 CCTACAGAAAGCTTCAAGAGTCTACCTCCCTATCCAGACATCCCC--GACTCTTCCC 928
DB      96839 CCTACAGAAAGCTTCAAGAGTCTACCTCCCTATCCAGACATCCCC--GACTCTTCCC 96780
QY      929 AMTTATTAAGACCCCCCTTCAACCAATGCTCCAAAAGAGATAGCAAAAAGGTAA 988
DB      96779 AAATATTAAGACCCCCCTTCAACCAATGCTCCAAAAGAGATAGCAAAAAGGTAA 96720
QY      989 CAGTAAACCAAGAGTGCATATTTCCCAATTATGACCCCTCCCAAGCAGTGGAGGA 1048
DB      96719 CAGTAAACCAAGAGTGCATATTTCCCAATTATGACCCCTCCCAAGCAGTGGAGGA 96664
QY      1049 GAGAAATCGGCCACAGCAGAGTGTGCTTTTCTCCCAAGCTTAAAGCAATTA 1108
DB      96663 GAGAAATCGGCCACAGCAGAGTGTGCTTTTCTCCCAAGCTTAAAGCAATTA 96608
QY      1109 AAACGACTTAAATTTCTCAATTAATCTGATGCTATATGTTGTTTAAAGGAT 1168
DB      96607 AAATGACTTAAATTTCTCAATTAATCTGATGCTATATGTTGTTTAAAGGAT 96548
QY      1169 TAGGCAATCTTTTATGATGATGAGATATATATGCTAGTAAATGACACTA 1228
DB      96547 TAGGCAATCTTTTATGATGATGAGATATATATGCTAGTAAATGACACTA 96489
```

```
QY      1229 ACCCAATATGAGAAATGCGCACATTAATGAGCCCTGAGGTGTTGGGATCTGGAT 1288
DB      96488 ACCCAATATGAGAAATGCGCGCAATATGAGCCCTGAGGTGTTGGGATCTGGAT 96429
QY      1289 CTCAGTACGCTTAATGATNGATGACCAACAGAAAGAAAGNATGATTTCCCAAGGC 1348
DB      96428 CTCAGTACGCTTAATGATNGATGACCAACAGAAAGAAAGNATGATTTCCCAAGGC 96372
QY      1349 CAGCAGCAGTTCCTCAGCTTASACCTCATTTGGGACAACAGAAATCAGTAATGAGGA 1408
DB      96371 CAGCAGCAGTTCCTCAGCTTASACCTCATTTGGGACAACAGAAATCAGTAATGAGGA 96316
QY      1409 TTGCTGCTGAGACATTTTCTTAATTGTGCTTASAGCTAAGAAATAGAGAA 1468
DB      96315 TTGCTGCTGAGACATTTTCTTAATTGTGCTTASAGCTAAGAAATAGAGAA-A 96257
QY      1469 ARTCTAATTAATCTCAATGATGTCACATTAACACAGGGGAGGAAAGAAATCTACT 1528
DB      96256 AGCTGTGAGTATTAATGATGTCACATTAACACAGGGGAGGAAAGAAATCTACT 96198
QY      1529 GCTTTTGGAGAGACTTAAGGAGGCAATTGAGAGCGTCTCT--CTGCACTGACT 1586
DB      96197 GCTTTTGGAGAGACTTAAGGAGGCAATTGAGAGCGTCTCTCTCTGCACTGACT 96138
QY      1587 CTCTGAAAGGCACTAATCTTAAGGGAATTAATCACTCAGTCAAGCTCAGACATTA 1646
DB      96137 CTCTGAAAGGCACTAATCTTAAGGGAATTAATCACTCAGTCAAGCTCAGACATTA 96078
QY      1647 -GAAAAAATCTCAAAAGTCTGCGCTAGCGCCGAGCAAACTTGAACCTTATGAAT 1705
DB      96077 GAAAAAATCTCAAAAGTCTGCGCTAGCGCCGAGCAAACTTGAACCTTATGAAT 96018
QY      1706 TGCGCACTGCGTTTATTAATGATGATGAGAGAGAGAGCGGAAACGAAACCGG 1765
DB      96017 TGCGCACTGCGTTTATTAATGATGATGAGAGAGAGAGCGGAAACGAAACCGG 95958
QY      1766 ATTAAAAAAGGCAACCGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1825
DB      95957 A-TAAAAAAGGCAACCGCTTATGATGATGATGATGATGATGATGATGATGATGAT 95899
QY      1826 GAAAAAGGAAAGCTGCGCAATTTGAATGCTTAATGAGGCTTCTTCAAGCGGTCTA 1885
DB      95898 GAAAAAGGAAAGCTGCGCAATTTGAATGCTTAATGAGGCTTCTTCAAGCGGTCTA 95839
QY      1886 CAAGGACACTTAAAAAAGATTGTCAGATGAAAGTAAGCCGCCCTTCTGTCATGCC 1945
DB      95838 CAAGGACACTTAAAAAAGATTGTCAGATGAAAGTAAGCCGCCCTTCTGTCATGCC 95779
QY      1946 TTATTTTCAAGGAATCACTGGAAGCCCACTGCCCGGAGGACAAAGGTCTTTTGA 2005
DB      95778 TTATTTTCAAGGAATCACTGGAAGCCCACTGCCCGGAGGACAAAGGTCTTTTGA 95719
QY      2006 GAAGCACTAATCAGATGATCAGACAGAGACTGAGGGTCTTGGGCAAGCGGCATCC 2065
DB      95718 GAAGCACTAATCAGATGATCAGACAGAGACTGAGGGTCTTGGGCAAGCGGCATCC 95659
QY      2066 CATGCAATCAGCTTCAAGAGCCCTGGGTATGCTTGAACATTTGAAGGCTCAGGAAGT 2122
DB      95658 CATGCAATCAGCTTCAAGAGCCCTGGGTATGCTTGAACATTTGAAGGCTCAGGAAGT 95599
QY      2123 TGTCTCTCGGACATGAGGCGGCTTCTTAATCTTAATCTTCTGCTCCGAGCAATGTC 2182
DB      95598 TGTCTCTCGGACATGAGGCGGCTTCTTAATCTTAATCTTCTGCTCCGAGCAATGTC 95539
QY      2183 TCCAGATCTGCTCACTATTTCTGAGGGGCTCCTTAAGACGAGTCACTAGTACTTTC 2242
DB      95538 TCCAGATCTGCTCACTATTTCTGAGGGGCTCCTTAAGACGAGTCACTAGTACTTTC 95483
QY      2243 CCAGCCACTAATGATTAAGCTGGGAGCTTATCTTTTCACTAGTCTTTCTTAATTATGC 2302
DB      95482 CCAGCCACTAATGATTAAGCTGGGAGCTTATCTTTTCACTAGTCTTTCTTAATTATGC 95424
```



2303 TTGAAAGCCCACTACCTTGTAGGAGAGACATTTCTAGCAAAAAGGAGGCAATTATAC 2362  
95423 CTGAAGGCCAATCCTTGTAGGAGAGACATTTCTAGCAAAAAGGAGGCAATTATAC 95364  
2263 ACCCTGAACATAGGAGAGAGAACCCCTTTGTGTCCTTGTGAGAGAGAAATTAAT 2422  
95363 ACCCTGAACATAGGAGAGAGAACCCCTTTGTGTCCTTGTGAGAGAGAAATTAAT 95309  
2423 CCTGAAGCTTGGGCAAGAGAGACATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2482  
95308 CTGAAGCTTGGGCAAG 95250  
2483 AGTTAACTAAAGAGATTCACCTTCTTCCCTAACAGAGAGAGAGAGAGAGAGAGAG 2542  
95249 AGTTAACTAAAGAGATTCACCTTCTTCCCTAACAGAGAGAGAGAGAGAGAGAGAG 95190  
2543 GGCCCAACAGAGATTCACCAAGATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2602  
95189 GGCTCAACAGAGATTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 95133  
2603 ATGATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2662  
95132 ATGATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 95080  
2663 AGTGAAGGTTAGTGAAGATTCAGATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAG 2722  
95079 AGTGAAGGTTAGTGAAGATTCAGATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAG 95022  
2723 AGCTGATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 2782  
95021 AGCTGATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 94962  
2783 AATCCGAG 2842  
94961 AATCCGAG 94902  
2843 GTTGAAGGTTAGTGAAGATTCAGATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAG 2902  
94901 GTTGAAGGTTAGTGAAGATTCAGATTAATCAATGAGAGAGAGAGAGAGAGAGAGAGAG 94842  
2903 AGGTTGAAGGTTAGTGAAGATTCAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAG 2962  
94841 AGGTTGAAGGTTAGTGAAGATTCAGATTAATGAGAGAGAGAGAGAGAGAGAGAGAG 94782  
2963 ATGATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 3020  
94781 ATGATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 94722  
3021 TCAGAAAGCTTGTGATTCAG 3080  
94721 TCAGAAAGCTTGTGATTCAG 94662  
3081 TACAGGTTTCCAAAG 3140  
94661 TACAGGTTTCCAAAG 94602  
3141 AATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3200  
94601 AATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94556  
3201 CCTGATTCAGAAAGCTTAAAG 3260  
94555 CCTGATTCAGAAAGCTTAAAG 94496  
3261 AATGATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 3320  
94495 AATGATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 94437  
3321 TCAGAAAGCTTGTGATTCAG 3380  
94436 TCAGAAAGCTTGTGATTCAG 94377  
3381 CCTTAAAG 3440

94376 CTAAAG 94319  
3441 TTTTAAATTCAG 3500  
94318 TTTTAAATTCAG 94264  
3501 GATGAGCTTCAACAG 3560  
94263 GATGAGCTTCAACAG 94204  
3561 GATGAGCTTCAACAG 3620  
94203 GATGAGCTTCAACAG 94144  
3621 AATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3680  
94143 AATTAATCTCCGCAATATTCGTAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94088  
3681 TAAAG 3711  
94087 TAAAG 94057

RESULT 8  
US-09-949-016-5640  
Sequence 5640, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949, 016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5640  
LENGTH: 2763  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-5640  
Query Match 29.1%; Score 2204; DB 4; Length 2763;  
Best Local Similarity 98.4%; Pred. No. 0;  
Matches 2194; Conservative 35; Mismatches 0; Indels 1; Gaps 1;  
5352 GATACATCTGAG 5411  
535 GATACATCTGAG 594  
5412 GCTGAG 5471  
595 GCTGAG 654  
5472 CTTGAAG 5531  
655 CTTGAAG 714  
5532 CTTGAAG 5591  
715 CTTGAAG 773  
5592 TTAATCAATTTTCTCTTAAATTTTCTCTTAAATTTTCTCTTAAATTTTCTCTTAAATTTT 5651  
774 TTAATCAATTTTCTCTTAAATTTTCTCTTAAATTTTCTCTTAAATTTTCTCTTAAATTTT 833

QY 5652 ATGCCGCTGATATGACAGTAGCTCCCTTACGMAAGTTTCTATGAGAAATGCAGCCTCC 5711  
DB ATCCCGCTGATATGACAGTAGCTCCCTTACGMAAGTTTCTATGAGAAATGCAGCCTCC 893  
QY 5712 CGGAAATATGATGCCCCCATCGTATAGAGCTTTTAAAGGAAACCCCACTTCACTGC 5771  
DB 894 CGGAAATATGATGCCCCCATCGTATAGAGCTTTTAAAGGAAACCCCACTTCACTGC 953  
QY 5772 CCAACCCATATGCCCCGCAACTGCTATACCTTGCACACTTTCATGATGCAATGCAATAC 5831  
DB 954 CCAACCCATATGCCCCGCAACTGCTATACCTTGCACACTTTCATGATGCAATGCAATAC 1013  
QY 5832 TCATATATGACAGGAAATATGATTAATCCAGTTGCTCCGAGAGACCTTGGAGCACTGT 5891  
DB 1014 TCATATATGACAGGAAATATGATTAATCCAGTTGCTCCGAGAGACCTTGGAGCACTGT 1073  
QY 5892 CTGTTGACCTTACTTCACTCCCAACTGGATATGCTGATGAGGAGTGAAGTCAAGTCAAGC 5951  
DB 1074 CTGTTGACCTTACTTCACTCCCAACTGGATATGCTGATGAGGAGTGAAGTCAAGTCAAGC 1133  
QY 5952 AAGAGAAAAATCTGTAAGAAAGATATCTCCCACTGACGSGGTATGATGACCTGTAG 6011  
DB 1134 AAGAGAAAAATCTGTAAGAAAGATATCTCCCACTGACGSGGTATGATGACCTGTAG 1193  
QY 6012 CCCCCAAGAGACTAGTCTCTCAAAATCATGAAACCTCCGTAACCTATCTCCCT 6071  
DB 1194 CCCCCAAGAGACTAGTCTCTCAAAATCATGAAACCTCCGTAACCTATCTCCCT 1253  
QY 6072 GGTAAAGCTATTTTAAATGACACCTCACTGGGCTCATGAGGTCTCGGCCCAAAACCTTAC 6131  
DB 1254 GGTAAAGCTATTTTAAATGACACCTCACTGGGCTCATGAGGTCTCGGCCCAAAACCTTAC 1313  
QY 6132 TAACTGTTGATATGCTCCCTCCCTGAACTTCAGCCATATGTTCAATCCCTGTAACCTGA 6191  
DB 1314 TAACTGTTGATATGCTCCCTCCCTGAACTTCAGCCATATGTTCAATCCCTGTAACCTGA 1373  
QY 6192 ACAATGGAACAATCTTCAAGACAGAAATTAACAACAATCTCCCTTTTAAAGAACTCTTGT 6251  
DB 1374 ACAATGGAACAATCTTCAAGACAGAAATTAACAACAATCTCCCTTTTAAAGAACTCTTGT 1433  
QY 6252 TTCCAAATGGAATTAATCCCACTGCTGAACTCACTGCTGTAATATTTAGCAATATATAC 6311  
DB 1434 TTCCAAATGGAATTAATCCCACTGCTGAACTCACTGCTGTAATATTTAGCAATATATAC 1493  
QY 6312 ATACACAACAATCTCCCAATGATCATGAGTGGGTAATCTCTCCACACAATATGCTGCT 6371  
DB 1494 ATACACAACAATCTCCCAATGATCATGAGTGGGTAATCTCTCCACACAATATGCTGCT 1553  
QY 6372 ACCCTCAGGAATATTTTGTCTGTGTGTAACCTGACCTATGCTGTTTGAATGGCTTTC 6431  
DB 1554 ACCCTCAGGAATATTTTGTCTGTGTGTAACCTGACCTATGCTGTTTGAATGGCTTTC 1613  
QY 6432 AGAATCTATGCTCTCCCTCATATTTAGTCCCTCCATGACCTATGCACTGAAACAAGA 6491  
DB 1614 AGAATCTATGCTCTCCCTCATATTTAGTCCCTCCATGACCTATGCACTGAAACAAGA 1673  
QY 6492 TTTATACAGTATGATCATATCTAAGCCCGGCAACAATAAGATACCATTTCTCTTTTGT 6551  
DB 1674 TTTATACAGTATGATCATATCTAAGCCCGGCAACAATAAGATACCATTTCTCTTTTGT 1733  
QY 6552 TATAGAGACGAGAGTGTAGGTGCACTAGGTACTGGCACTTGGCGGTATACAACTCTAC 6611  
DB 1734 TATAGAGACGAGAGTGTAGGTGCACTAGGTACTGGCACTTGGCGGTATACAACTCTAC 1793  
QY 6612 TCAGTTCTACTACAACTATCTCAAGAACTTAATGGGACATGGAACGGGTGCGCGATTC 6671  
DB 1794 TCAGTTCTACTACAACTATCTCAAGAACTTAATGGGACATGGAACGGGTGCGCGATTC 1853  
QY 6672 CTTGGTCACTTGGCAAGATCAACTTAATCTCCCTAGCAGCAGTATGCTTCTTGAATCGAG 6731  
DB 1854 CTTGGTCACTTGGCAAGATCAACTTAATCTCCCTAGCAGCAGTATGCTTCTTGAATCGAG 1913  
QY 6732 AGCTTTAGACTTGTCTAACCCGCTGABAAGGGGGAACCTGTTATTTTAAAGGGAAGATG 6791

DB 1914 AGCTTTAGACTTGTCTAACCCGCTGABAAGGGGGAACCTGTTATTTTAAAGGGAAGATG 1973  
QY 6792 CTGTTATTAATGTTAATCAATCCGGAATGCTCACTGAGAAAGTTAAGAAATTSAGATCG 6851  
DB 1974 CTGTTATTAATGTTAATCAATCCGGAATGCTCACTGAGAAAGTTAAGAAATTSAGATCG 2033  
QY 6852 AATACAACTTAAGAGAGAGAGCTTGAACCTGAAACCTGAAACCTGGGGCTCTCAAGCCATG 6911  
DB 2034 AATACAACTTAAGAGAGAGAGCTTGAACCTGAAACCTGAAACCTGGGGCTCTCAAGCCATG 2093  
QY 6912 GATGCCCTGGAATCTCCCTCTTGAAGACTCTGAGAGCTATTAATATGCTACTCTCT 6971  
DB 2094 GATGCCCTGGAATCTCCCTCTTGAAGACTCTGAGAGCTATTAATATGCTACTCTCT 2153  
QY 6972 TGAACCTGTATCTTAACTCTCTGTTAACTTGTCTCTTCCAGAAATCGAATCTGTAA 7031  
DB 2154 TGAACCTGTATCTTAACTCTCTGTTAACTTGTCTCTTCCAGAAATCGAATCTGTAA 2213  
QY 7032 ACTACAAATGAGACCCCAAGATGCAATGCAAGACTAAGTCTACGAGACCCCTGAGCCG 7091  
DB 2214 ACTACAAATGAGACCCCAAGATGCAATGCAAGACTAAGTCTACGAGACCCCTGAGCCG 2273  
QY 7092 GCTGTAGAGCCCAAGATGCTGATGTTAATGACATCAAGGACCCCTCTGAGAAATCTC 7151  
DB 2274 GCTGTAGAGCCCAAGATGCTGATGTTAATGACATCAAGGACCCCTCTGAGAAATCTC 2313  
QY 7152 AGCTGACAACTCTTACATGAGCCCAATTCAGAGAGAGAGCTTAAGCCGTGCTGAGCC 7211  
DB 2314 AGCTGACAACTCTTACATGAGCCCAATTCAGAGAGAGAGCTTAAGCCGTGCTGAGCC 2393  
QY 7212 AACCTCCCAAGAGACTTAAAGTTTCTGTTGAGATGAGGAGAGCTGAGAGACAGACTAG 7271  
DB 2394 AACCTCCCAAGAGACTTAAAGTTTCTGTTGAGATGAGGAGAGCTGAGAGACAGACTAG 2453  
QY 7272 CTGGAATTTCTGAGCTGATTAAGAAATCTTAAGCTTAAGTGGGAAAGGTGACCAATCTAC 7331  
DB 2454 CTGGAATTTCTGAGCTGATTAAGAAATCTTAAGCTTAAGTGGGAAAGGTGACCAATCTAC 2513  
QY 7332 CTTTAAACAGGGGCTTGCATTAAGTCAACCTGACCAATCAAGAGAGCTCACTAAAT 7391  
DB 2514 CTTTAAACAGGGGCTTGCATTAAGTCAACCTGACCAATCAAGAGAGCTCACTAAAT 2573  
QY 7392 GCTAATTAAGCAAGAGAGGTTAAGAAATGAGCCATGATATGTTGTTGAGAGAGACA 7451  
DB 2574 GCTAATTAAGCAAGAGAGGTTAAGAAATGAGCCATGATATGTTGTTGAGAGAGACA 2633  
QY 7452 GAGAGAGGAGCAATGATCGGATATTAACCAAGTTCGAGCCGCAACGGCAACCC 7511  
DB 2634 GAGAGAGGAGCAATGATCGGATATTAACCAAGTTCGAGCCGCAACGGCAACCC 2693  
QY 7512 TTTGGGTCCCTCCCTTGTATGAGAGCTGTTTCACTATTTCACTATTAATC 7571  
DB 2694 TTTGGGTCCCTCCCTTGTATGAGAGCTGTTTCACTATTTCACTATTAATC 2753  
QY 7572 TTGCACTGC 7581  
DB 2754 TTGCACTGC 2763

RESULT 9  
US-09-573-080A-26  
Sequence 26, Application US/09573080A  
Patent No. 6828097  
GENERAL INFORMATION:  
APPLICANT: JOAN, KNOLL  
APPLICANT: ROGAN, PETER  
TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENER  
FILE REFERENCE: 30307  
CURRENT APPLICATION NUMBER: US/09/573, 080A  
CURRENT FILING DATE: 2000-05-16  
NUMBER OF SEQ ID NOS: 479  
SOFTWARE: PatentIn version 3.0

```
seq id no 26
: LENGTH: 8399
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: repeat_region
: LOCATION: (1)..(839)
: OTHER INFORMATION: hepv9
: PUBLICATION INFORMATION:
: AUTHORS: Jurka, J; Malichiewicz, J; Milosavljevic, A
: TITLE: Prototypic sequences for human repetitive DNA
: JOURNAL: Journal of Molecular Evolution
: VOLUME: 35
: ISSUE: 4
: PAGES: 286-291
: DATE: 1992-10-
: DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)
: DATABASE ENTRY DATE:
: DATABASE ENTRY DATE: 1996-01-26
: US-09-573-080A-26
```

```
Query Match 28.6%; Score 2168; DB 4; Length 8399;
Best Local Similarity 67.6%; Pred. No. 0;
Matches 3334; Conservative 143; Mismatches 1354; Indels 103; Gaps 34;
```

```
Qy 605 TGATGGGAAAGCTTCCCGCAAGCAAAACGCCCTAAGACGTATCTGGABAATTGG 664
Db 1276 TGATAGGAAACATCAGGACATCAACACGCTCAACCTTGAAATGTATCTTAAGCATTGG 1335
Qy 665 AMCAATTTGACCTCTGACACTAAGAAAGAAAGCACTTATTTCTTCTGCAAGTCCGCT 724
Db 1336 ACCAATTTGACCCCAAAACCTGAAAGAGGCACTATTTTCTGCACTACAGCT 1395
Qy 725 GGCACCT-----CTGAGGAAAGTAAATTAT 752
Db 1396 GACCTTAATATCTCTCTGATGGGAAAAATGGCCACTGAGGAAAGTAAATTATCA 1455
Qy 753 ACACCATCTTACAGCTAGACATCTTTTGT---AGAAAGGCAAAATGAGTAAAGCCAT 809
Db 1456 ATACTATCTGCAAGTTCATCTTTCTTAAGAGGAAAGTAAATGAGTAAATACAT 1515
Qy 810 AAG---TACAACTTTCTTTTCAATTAAGACAACTCAATATATGTAAGAAAGTGAT 865
Db 1516 AAGGTATTAACAAGCTTTCTTTTCACTGAAGGAAATATACAACTATGCAAAAGCTTGCAAT 1575
Qy 866 TTATGCCCTAAGAGGCTTCAAGTCTACCTCCCTATCCAGATCCCGAGCTCCCTC 925
Db 1576 TTACATCCCAAGAGGACCTTTCAGCTTACCCCAATCTAGGCTCCCTAGAGCTCTC 1635
Qy 926 C--CCAAATTAAGAGCCCTTCAACCAATATGTCCAAAAGAGATAGACAAAG 983
Db 1636 CTCTCTATTAATGCGAGGCTCTCTCCATCTCCCTGCCAGAGAAATTAAGCAAGAA 1695
Qy 984 GTAACAGTGAACCAAGAGTGCATATATCCCAATATGACCCCTCCCAAGAGTGG 1043
Db 1696 ATCTCAAGAGACCAAAACCCCTGAGCTATGCTATGTCCTT-CAAGCTGTAGG 1754
Qy 1044 AGGAAGAAATTTGGCCCAAGCAGATGCTGTTTCTCCCAAGCTTAAAGCA 1103
Db 1755 GGGTGGAGAAATTTGGCCCAAGCAGATGCTGTTTCTCCCTGCTGATTTAAAGCA 1814
Qy 1104 AATTAATAACAGACTTAAGTAAATCTCAGATTAATCTGATGCTATATGTTTACA 1163
Db 1815 GATCAAGGACAGCTGGGGAAGTTTCAAGATGATCTGATAGGACATAGATGCTCCA 1874
Qy 1164 AGGCTTGAAGCAATTTCTTGAATCTGACATGAGAGATATATATGCTACTGCTAAATCA 1223
Db 1875 GGGCTAGGGCAAACTTCGATCTCACTTGAAGAGATGT-CATGCTATTTGATAGTCAA 1933
Qy 1224 CACTTAACCCCAATGAGAGAGTGCACATTAATCTGAGCTGAGAGTTTGGCAATCTCT 1283
Db 1934 CCGTGGCTTTTAATGAAAAAGATGTGCTGATGACGCTGAGAGTTTGGAGATTA 1993
```

```
Qy 1284 GGTATCTCAGTCAAGTCAATGATANGATGACAAAGAAAGAAAGAAATGATTCCTCA 1343
Db 1994 GGTATCTTATGCAAGTAAATG---ATGAATGACAGCGAAGAAAGATAAATTTCCCTA 2050
Qy 1344 CAGGCGACAGAGAGTTCCAGTCTASACCTCTATGGGACACAGAAATAGTAACTG 1403
Db 2051 CTGGCCAGCAAGCATCCCAATGATGATCCCACTGGGACCTC---AACTCAGATCAT 2106
Qy 1404 GGAGATTGGTCTGACACATTTGCTAATCTGTGTGCTASAAAGGACTAAGGAAACTASG 1463
Db 2107 GGGACCTGAGTCTTAACATCTGTGACCTGTGTGTAAGGACTAAGGAAATTTAGG 2166
Qy 1464 AAGAAATCTAAGATTAATCAATGATGTCACCATTAACAGAGGAAAGGAAAGAAATC 1523
Db 2167 AA-AAAGCCATGATTTATTCATGATGTCACCATTAATCA-AGAAAGGAAAGAAATC 2224
Qy 1524 CTACTGCTTTCTGAGAGACTTAAGGAGCTTGAAGAAAGCTGCTCTTCTCACTG 1583
Db 2225 CTCTGCTCTTCTCAGTGGCTTCAAGAGACCTTAAGAAATATATCTCACTGACCTG 2284
Qy 1584 ACTCTTGAAGGCAACTAATCTTAAGGCTAAGTTATCACTCAGTCCGTCAGACA 1643
Db 2285 AATCACTGAGGCTCAATTAATCTTAAGATTAAGTTATTAATCACTCAGACAGATA 2344
Qy 1644 TTAG-AAAAAATTCAAAAGTCTGCTGAGCCCGGAGCAAACTTAAGAAACCTATTGA 1702
Db 2345 TCAGTAAAGGCTCCAAAGCAAGCCCTGGGCTTGAACAAATATGAGGCAATTAATTA 2404
Qy 1703 ACTTGGCACTGTGGTTTTTATTAATGAGATCAAGAGAGAGCAGGCGAAGCAAAAC 1762
Db 2405 ACCTGGCAACCGTGTGTTTATTAATGAGACCAAGAGAGCAAGGCCAA---AGGAA 2460
Qy 1763 GGGATTAATAAAAGGCAACCGCTTATGATGATGACCTCAGGCAAGTGAAGCTTGAAGC 1822
Db 2461 AGAGATCAAGAAAGGCGGAGGCTTATGATGATGAGCTTCAAGCAAACTTTGATGT 2520
Qy 1823 TCTGAAAGAGGAAAGCTGGGCAATTAATGATGCTTAATAGGCTTCTTCACTGCGGT 1882
Db 2521 TCAGAGAGGACAGAAATGAGAGGCAATCAAGCTGTAGGCTTATTAATCACTGTGT 2580
Qy 1883 CTACAGGACATTTTAAAGATGTCGCAAGTAAAGTAAAGCCCTTCTGCTCAATG 1942
Db 2581 TTACAAAGACATTTTAAAGATGTCGCAAGTAAAGCAAGCTTCTTATCTGTGT 2640
Qy 1943 CCTTATTTCAAGGAAATCACTGGAAGGCCCACTGCCCCAGGAGCAAAAGTCTTTGAG 2002
Db 2641 CTACTATGCGAGGCAATCACTGGAAGTGAATGCTGCCCCAGAGCAAAAGTTCTAGG 2700
Qy 2003 TCAGAGGCACTAAGCAGATGATCAAGCAGAGACTGAGAGGTGCTGGGCGAAGCCCA 2062
Db 2701 TTGGAAGCCCCCAACCTGATGATCAACAAGGACTGAGAGGTGCTGCAAGGCCCA 2760
Qy 2063 TCCCATGCGATCAACCTCAAGAGCCCTGGTATGCTTGAACATGAGGCGCAAGAA 2119
Db 2761 GCTCATGCTATCAACCTCACTGAGGCCCAAGATGTTAACATTAAGAGGCGCAAGAAAT 2820
Qy 2120 GGTGTCTCTGGAACCTGAGTGGGCTTCTTATGCTTACTTCTTCTGCGCAACTG 2179
Db 2821 GACTTCTCTTGAAGCTGCAAGTCTTCTGAGTGTATCTCTTCTGAGATGACTG 2880
Qy 2180 TCTTCAGATCTGCTCACTATTTGAGAGGCTGCTAAGAGCGGAGTCACTAGTACTT 2239
Db 2881 TCTCAAGAGTCTGTACATCC---GAGGAATCTGGGCGAGCTGTACAGAGT-CTTT 2296
Qy 2240 TTTCCAGCCCTAAGTATTAATCTGAGGAGCTTATCTTTTCAATGATTTTCAATTA 2299
Db 2937 CTCCCACTCTCAAGTGT-AAATGGAGACTTGTGCTTCTGTATGCTTCTTGTGTTA 2995
Qy 2300 TGCCTGAAGCCCACTAAGCTTGTAGGAGAGCAATTTAGCAAAAGAGGCGCAATTA 2359
Db 2996 TGCTGAAGTCCCACTTATTAAGGAGATATAT--TAGCCAAAGCTGAGACTATTA 3053
```

OY		2360	TACACTGAACTAGAGAAGACAACCCTTTGTGTGCCCTTGTAAGGAAGATT	2419
Db		3054	TCATCATGAATATGGGGAACAAGTATCCCAATTTGTTG-CCCTTAATTGAGAGGAATT	3112
OY		2420	AATCGTAAGCTGGGCAACGAAGGACATAATGAGCAGAACAAAGATGCGCTGT	2479
Db		3113	AACCTGAAGCTGGGCATTGGAGACAAATTGGAMGGCAAA-NAITGCCACCACT	3171
OY		2480	TCAAATTAACTAAAGGATCCACTCTCTTCCCTCACCAAGGAGTAAACCCCTCAGAC	2539
Db		3172	CCAATATCAGGGCTAANAAGCTCCACACATCTTCTTATCAAGGCAATATCCTTAAGGCC	3231
OY		2540	CAGGCCCAACAGGATTCCAAAGATTGTTAAGACTTAANAAGCCCAAGCTTAGTAA	2599
Db		3232	TGAACCTCTTAAGGATTACAGGATATGTTAAACATTTAAAAGTCAAGGCTTAGAG	3291
OY		2600	ACCAATGCAATACCTCCCTGAGTAATTTCCGTAGTGATGAGAGGCAAGAAACCCAGTG	2659
Db		3292	GAAATGACGACTCCCTGACACCCC-----NATCTAGAGTACAAABACCAATG	3344
OY		2660	GACATGAGGGTTATGTCAGAGTCTCAGATTATCAATGAGGCGCTGTCTTTTATA	2719
Db		3345	GTCATCTGA-GACTAGTCAGATCTTAGACTATCAAT-GAGCAGTAATTCCTTATA	3402
OY		2720	CCGAGCTGTACCTAGCCCCCTTATCTGTGMYTTCCCAATACCAAGAGCAGATGTT	2779
Db		3403	TCGATTTGATCCCAACCCCTATACCTGTCTTCAATPACAGAGGAGCAGATGTT	3462
OY		2780	TACASTCTGGAACCTTMAAGATGCTTCTTGCAATCCCTGTAATCCTGACTCAATT	2839
Db		3463	CATGTTCTGGACCTCAAGGAGCTTCTTCTGTGTCC-CGCACTCTGACTCCAGTT	3521
OY		2840	CTGTGTTGCTTGAAGTACTTCAACCCACATCTCACTCAGCTGACTRTTTTACC	2899
Db		3522	TCTGTTGCTTGAAGATCCACAGACCAACGTCCTCAATCATBGAATGCTTGCC	3581
OY		2900	CCNAGGTTCAAGGATAGYACCCCATCTATTGGCCAGGCAATTAGCCCAAGCTTAGCYCA	2959
Db		3582	ACAAAGGTTTAGGATAGCCCTCACCTGTTGTGACGGCACTGGCCCAAGTATAGGCCA	3641
OY		2960	KTYVTCATACCTGACACTCTTGTCTCTGAKGTGATGATTTACTTTTREGYGCYR	3019
Db		3642	CTTCTCAAGTCCAGGCACTTGTGCTTCAAGTATGATGATGYTAACTTTGGCTACAG	3701
OY		3020	TTCAAGAAACCTGTGCATCAAGCAACCCACACRCTTMAATTTCTCGGACCTGGG	3079
Db		3702	TTCAAGAGCTCATGCAAGCAAGCTTATATCTTTGAATCTTGAATCTTAATCAAGS	3761
OY		3080	CTACAWGTTTCCAAACSARAGCTCARCTGTCTCACAGCAGTTAAATCTTAGGCT	3139
Db		3762	GRCAAGGATCTAGSTCAAAAGTGAGCTTTGCTTAAGCAGGCTAATATCTAGGCT	3821
OY		3140	AAPATTATCCAAAGCACARGGCCTCATGAGAAATRVATCCAGCTTATCTGCTTA	3199
Db		3822	AATCTTAGCCAAAGGACCAAGGCTCTCAGCAAGGATGAATACGCTTAATCTGCTTA	3881
OY		3200	TCTCATCTCAAAACCTTAAAGCACTAAGRBRTTCCCTGGCRPAAYAGTTTGTGCG	3259
Db		3882	TCTTTCCTTAAGACTTAAAACAGTTGACGAGGTTCTTGGAAATCACCGGCTTTTGCG	3941
OY		3260	AAWATGATTCGCCAGGTGCGRAAATAGCCAGGYCATTAATACASTAATTAAGAA	3319
Db		3942	ACTATIGA-TCCCTGATATCAGTGAGATGTAGGCGCTTCATCTCTAATCAAGGA	4000
Db		4001	CCAGAGGCAATAATCTTATCTAGTAGATGTATACAGGGGCAAGAACAGCTTCAAA	4060
OY		3320	CTCAGAAACCAATACCAATTAATTAAGTGAIVAMCTGAAYMBAGTGCTTCCAGS	3379
OY		3380	CCCCTAAGAAGCTTAAACCAAGYCCCAAGTTAAATYTGGCAORGGCAAGACTT	3439
Db		4061	CCTTAAAGAGGCGCTGTG---CAATTCACCTTTAAGCCTTCCAC-----NACTT	4111
OY		3440	TTSTTATATFTCACAGAAAAAACGRAAVAGCTTRBGATCTCTTACACAGTCCAG	3499

Db	4112	CTCTTTATATGTCACG-AGAGAGCAGGATACCTTGGAGTCTTATCTCAGACTCTTG	4170
Qy	3500	GGATGACCTTGCAACCCYRTGGCRVACCTGASTAGGAAATGATGTAGTGGCAAGGTT	3559
Db	4171	GGACAAACCCCAACCAACAGTGGCATTACTAATGAGAAATTTGATGTAGCAAAAGGCT	4230
Qy	3560	GRCCTCATTTGTTATVGGGTATGTGTGGCACTTACAGTATKATATCTGAAGCAAGTTAAA	3619
Db	4231	GGCTCACTCTTTAAAGGTATGGTTCAGCACTGGCCATCTTCATGTCCAGAGCATTA	4290
Qy	3620	TAATACAGGAGAGATCTTACTGTGTGACATCTCATGAKGTGAATRGCATACTG	3679
Db	4291	TAATACAGGAA-AGATCTCACTGTCTGACATCACTATGTATTAATGGCATCTAGGTG	4349
Qy	3680	CTAAAGAGACTTGTGGCTGTCAACAATCTTTACTTAATFTCAGGCTCTATTACTTG	3739
Db	4350	CCAAAGAAATTTATGCTATTCAGACTACCACTACTATGATACAGGCACTACTCTCTTG	4409
Qy	3740	AARGGCCAGTGTGCACTGTGACCTTGTCACCTCTAACCCAGTNCATTTCTTCAG	3799
Db	4410	AGGACCAAGTCTCAAAATCTACAGTGTGGCCCTCAACCTTGCCACTTCTCTCCAG	4469
Qy	3800	ACATATGAAGAAAGATBARATATACGTGTCAACAATTAATTTCTCAACCTATGCCATC	3859
Db	4470	AGAAATGGGAACCAATTTAGACATGATCTGCCAATAATACATGTCAGACTTATGCCACC	4529
Qy	3860	GAGGGGACCTTTTAGAGTTCCTTTGACTGATCCYGACCTTCAACTTTATCTGATGA	3919
Db	4530	AAGATGATCTCTTAGAAGTCCCTTACTATCTCATCTTGACCTT-AACTATATACCGATGA	4588
Qy	3920	AGTTCCTTTGTGAAAAAGACTCCAAAAGYGGGGATGCAAGTGTCACTGATTAATGGA	3979
Db	4589	AGTTCATTTGCAAAAATGGGATATGAAGGGCAAGTTGTACATATGTTATGTATCT	4648
Qy	3980	ATATTTAAGATATCCCTCACTCCAGGAACATAGTCTATGCTRGCAAGACTAATAGCC	4039
Db	4649	GTACCTGAATTAAGCTCTTCCCCAGGGACCAAGCCAGTTTAACGAATAGTGGCA	4708
Qy	4040	YTCAATYKGGCACTAGATTTAGAGAGRAAAAAGGYAATAATATATACAGACTTAR	4099
Db	4709	CTTACCTGAGCTTGAATCTGGAAAAGGAAAAAAGATTAATGTGATATACATAGCAAG	4768
Qy	4100	TATGCTTACCTATGTCNTCCATGCCCCATGATGACATATGSRAGAAAGGAAATTTCTTA	4159
Db	4769	TATGCTTATCTAATCCCTCATGCCCCATGTGCAATATGAAAGAAAGGAGTTCTTA	4828
Qy	4160	TCYAGRRAACCTATTCAMACATCAGAAAGCCATTGGARATATTAATATYTGCGWTACAG	4219
Db	4829	TCCGGGGTACCCCCATTTAAATGCGACAGAGATTTATGGATTATTTGCACTGCA	4888
Qy	4220	AAACTBARAGAGTGTGMAGTCTTCACTGTGTGGGTATCCATCAAAGAAAGRAAGGAA	4279
Db	4889	AAAGCCAGAGAGGTGTGACGTCTTACATCTGCCAAAGCATCAGAAAGTGAAAGGAAAG	4948
Qy	4280	ATASAAAGGAATYTGCAAGCAKATATTTGAACMAAAAGCTGCAAGCAGGACCTCTCA	4339
Db	4949	GCAGAAAGAAACCGTCAGGCGAGATGTGAGGCCAAAAATTCGTCCAGGTGATCTCCCA	5008
Qy	4340	TTAGAAATGCTTATTTAACTTCCCTTATGATTAAGGTATCCCTTCCGGGAAACCAAGCC	4399
Db	5009	TTAGAAATTACTATGAAAGGACCTTTGAAATGAAACAAACCTTCCAAAGATTAAGCC	5068
Qy	4400	CAGTACTCAGCAGGAAACAGAAATGGGAAACCTCAGAGG---CAGTTTCTCCCTCG	4456
Db	5069	CAGTATTTCCCAATGAAACGAATGGGACTCAGTGGGGGCAATGTTTCTCCCTCA	5128
Qy	4457	GGACGGTAGCCACTGAAGAGGAAAAATCTTTTGGCTGCAATATTCATATGAAATTA	4516
Db	5129	GGGTGTTAAAGCAAAAGGAAAGGATCTTATATCCGAAAGCAGCAAGTGGAAATA	5188
Qy	4517	CTTAAACCCCTTCACTAACTTTCACTTAGGATCGATACCAACCATCAATGGCCAA	4576







```

1 PRIOR FILING DATE: 2000-10-03
2 PRIOR APPLICATION NUMBER: 60/331,498
3 PRIOR FILING DATE: 2000-09-08
4 NUMBER OF SEQ ID NOS: 207012
5 SOFTWARE: FastSeq for Windows Version 4.0.
6 SEQ ID NO 13840
7 LENGTH: 276687
8 TYPE: DNA
9 ORGANISM: Human
10 FEATURE:
11 NAME/KEY: misc_feature
12 LOCATION: (1)...(276687)
13 OTHER INFORMATION: n = A,T,C or G
14 US-09-949-016-13840

```

Query Match	19.9%	Score 1507.4	DB 4	Length 276687
Best Local Similarity	68.3%	Pred. No. 0		
Matches 2889	Conservative	94	Mismatches 914	Indels 52
				Gaps 24

QY	697	CGACTTATATCTTCTGTGAGTGGCCGCTGGCAGCTCTGAGGAAAGTTAAATATTAACG	756
Dp	27184	CCAATATTTTCTCTGTATGGGAAAATATGGCCACTGAGGAAAGTCAAAATTACATAC	27243
QY	757	CATCTTACAGCTAGACYCTTTTGTAGAA--AAGGCAATGAGTGAAGTGGCATTAAGT	813
Dp	27244	TATCTCAGAGCTTACCTTTTCTGTAGAGGAAAGCAATGAGTGTGAATCCTTATGT	27303
QY	814	ACAAACCTTTCTTTTCATTTAAGAGCACTCAATATTATGTAAAGAAGTGTATTAAGCC	873
Dp	27304	CCAAGCTTTCTTTTCATTGAGGGAATACACAGCTATGCAAAAGCTTGCAATTTATCATCC	27363
QY	874	TACAGGAAGCTTCAGAGTCTACTCTCCCTATCCAGACATGCC--GACCTCTTCCCCAM	930
Dp	27364	CACAGGAGACCTCTCAGCTCACCCTCATCTCTAGCTCTCTATATAGCTCCCTTCTCAT	27422
QY	931	TTAATAAGAACCCCTCTTCAACCCCAATGGTCCAAAAGAGATAGACAAAAGGTAAACA	990
Dp	27424	GAATGATTAATCTCTCTTAATCTCCCCCGCCGAAAGAAATAGCAAAAGAAATCTCCA	27483
QY	991	GTGAACCAAAAGTGGCCAAATATTTCCCAATTAATGACCTTCCCAAGCAGTGGGAGAA	1050
Dp	27484	AAGGACCACAAAACCTCTCGGCTATTGTATTGTCCCTT-TCAAGCTGTAGGGGAGGG	27542
QY	1051	GAATTTGGCCCAAGCAGAGTGCATGTGCTTTTYYCTCCCAACTTAAGCAATAAATA	1110
Dp	27543	GAATTTGGCCCAACCCGGGTACATGTCCCTCTCTCTCTAATTTAAAGCAATCAAG	27602
QY	1111	ACAGACTTAGGTAAATCTCAGATTAACCTGATGGCTAATGTGTGTTTACAAAGGTTA	1170
Dp	27603	GCAGACCTGGGAAAGTTTCAGATGATCTGTATGTGTATGATACATAGTGTCTTCAAGGTCTA	27662
QY	1171	GGACAATTTCTTGTATCTGCATGGAGATATATATGTCACTCTAAATCAGACACTTAAC	1230
Dp	27663	GGGCAACCTTGCACCTCGCTTGGAGATGT-CATCTCATGTGAGATCAAACTTTGGC	27722
QY	1231	CCCAATGAGAAAGTGCACACTAATCTGCAGCTGAGGTTTGGCGATCTCTGTATCT	1290
Dp	27722	CTTAAATGTAAAGAAATCGGCTTTTACGTGACGCCCAAGATTGTGGATATCCTGTATCT	27781
QY	1291	CAGTCAGGTCAATGATATANGATATGACAACAGAAAGAAAGAAATGATTTCCCAACAGCCA	1350
Dp	27782	TAGTCAAGTAATNT--ATAGAAATGACAGCCGAAAGAAAGGCAAAATTCCTTACCGGTTA	27838
QY	1351	GCARGCAGTTCCTCAGTCTTAAACCTCATTTGGGGACACAGAAATCAAGTAAATGAGGATTT	1410
Dp	27839	GCAACCAATCCCACTATGTGATCCCCACTGGAC--TTGACTCAGATCATAGGGGACT	27894
QY	1411	GGTGTGCAGACATTTGTCTAACTTTGTGTCTAABAGACTTAAGAAAATCTAGAGAAAR	1470
Dp	27895	GGAGTCAATTAACATCTGTGCTGTATTTCTAGAAAGCACTAAGAGAAATTA-GAATAAAG	27953
QY	1471	TCTATGAATTACTCAATGATGTCCACTTAACAGGGGAAGGAAAGAAATCTTACTGC	1530

Db	27954	CCCATGAATTATTCATGATGATGCCCATPACTCA--GGAAAAGAAAGAAATCTTCTGCG	28012
Qy	1531	CTTTCGTGAGAGACTAAGGAGGCAATTGAGAAAGCGTGCCTCTCTGTCACTGTGACTCTTC	1590
Db	28013	CTTCTCTGAGCGGCTACGGGAGGCGCTTAAAGAAAATATATCTCCCTGTCACTCCGAATCACT	28072
Qy	1591	TGAAGGCCAATCTATCTTAAAGGCTAAAGTTTATCACTCACTGATCACTGTCAGACTTTGAAA	1650
Db	28073	TGAGGGCTCAATTATCTTAAAGATTAAGTTTATTTACCGATGACGCAAGATATCGAGAG	28132
Qy	1651	AAACTTCAAAAGCTGCGCCGTAGGCGCCGAGCAAAACTTGAAGAACTTATTTGAACTTGACA	1710
Db	28133	AAAGCTGAAAAGCAATTCAGGGCGCTGAAACAAAATCTGAGGCAATTATTAACCTGGCA	28192
Qy	1711	ACTCGGTTTTTTATATATAGATCAAGAGAGACAGGCGGAAACAGGACAAACGGGATTAA	1770
Db	28193	ACCTCGGTGTCTTATATATAGGAGCAAGAGAGAAACAGGCCAAAGAAAAGGAGATCAG	28252
Qy	1771	AAAAAAGGCCACCGCTTTTATGTCTATGACCTTCAGGCAAGTGAATTTGGAGGCTCTGGAA	1830
Db	28253	AGAA-AGGCGGACGCTTATGTCTATGCGCCCTCAGACAAACAAACCTTGGTGGTTCAAGAG	28311
Qy	1831	AGGAAAAAGCTGGGCAAAATTGAATGCTTATATAGGCGTTGCTCTCAAGTGGGCTACAAAG	1890
Db	28312	GACAGAAAATGAGGACGAGGCCAATCACTGCTGTAGGCTTTTATCACTGTGGTTTACTAGG	28371
Qy	1891	ACACTTTAAAAAGATTGTCCAATGAAGTAAGCCGCGCTTGTCTCAATGCGCCCTTAT	1950
Db	28372	ACACTTTTAAAAAGATTGTCCAATGAAGTAAGTAAGCCGCGCTTGTCTCAATGCGCCCTTAT	28431
Qy	1951	TCAAGGGAATCACTGGAAGGCCCATCTGCCCCAGGGGACAAAGCTCTTTGATTCAGAAC	2010
Db	28432	CCGAGGCAATCACTGGAAAGGTGACTGCCCCAGAGACAAATTCCTCGGGTCAAGAAC	28491
Qy	2011	CACCTAACAGATGATCATCAGACAGGACTGAGGGTCTCGGGGCAAGCGCCATCCCATGC	2070
Db	28492	CCCCAACAGATATATCAACACAGGACTGAGGGTCTCGGGGCAAGCGCCATCCCATGC	28551
Qy	2071	CATCACCTTCACAGAGCCCTGGGTATCTGTGACCATTGAAGGCGCAAGAA--GGTTGTCT	2127
Db	28552	CATCACCTTCACAGAGCCCTGGGTATCTGTGACCATTGAAGGCGCAAGAAATTTGACTTCT	28611
Qy	2128	CTTGGACACTGTGTGCGGTCTTCTTAATCTTACTTTCTGTCCCCGACAACTGTCTCCAG	2187
Db	28612	CTTGGACACTGTGTGCGGTCTTCTTAATCTTCTGTCTCGGTGACTGTCTCTCAGG	28671
Qy	2188	ATCTGTACATTTCTGAGGGGGGCTTAAGAAGGAGACTCATGATCTTATTTTCTCCAGC	2247
Db	28672	----TCTCTTACATCCAGAGAAATCTGGGACAGCTGTACAGGTA--TTTCTCCACC	28726
Qy	2248	CACCTAAGTTATGAATCGGGAGCTTATATCTTTTCAATCATGCTTTTCTAATATATGTGAA	2307
Db	28727	TCTCTAGTTGT-AATTGGAAGACTTTGCTTTTCAATCATGCTTTTCTAATATATGTGAA	28785
Qy	2308	AGCCCCACTACTTTGTTAGGAGACACTTTAGCAAAAGCAGGGGCAATTATCACTG	2367
Db	28786	AGTCCCACTCTTATTTAGGAGGAGTATATTAGCCAGGCTGTGAGCTTTTATCTACATG	28845
Qy	2368	AACATAGGAGAGAGAACCCCGTTGTGTGTCGCCGCTTGTGAGGAAGAAATTAATCTGA	2427
Db	28846	AATATGAGAAACAATTATCCCAATTTGTGT--CCCTCATTTGAGAGGAGATCAACCTCTGA	28904
Qy	2428	AGTCTGGGCAACGAAGGACAATATGAGAGAGCAAGATGCGCGTCTGTTCAGTTA	2487
Db	28905	AGTCTGGGCAATGGAAGGACAATTTGGAAG--GGCAAAAATATGCTGTGCCAATATCA	28963
Qy	2488	AACCTAAAGATTCACTTCTTTTCCCTACCAAGGCAATGCCCTCAGACCCCAAGGCC	2547
Db	28964	GGTTAAAAAGTCCACCACTTTTCTTATCAAAAGGCAATATCTTTAAGGCTGTGAAGCTC	29023
Qy	2548	AACAAGATTCCAAAGATTTGTAAGGCTTAAAGCCCAAGGCTTATGTAACCATGCA	2607
Db	29024	ATTAAGATTACGAATATTTGTTAAACTTTAAAGCTTCAGGCTTATGTAAGAAATGCA	29083

OY	2608	TAACTCCCTGACGTAATTTCCGTAGTGGATTGAGAGAGGACAGAAACCCAGTGGACAGGG	2667
Db	29084	GCACTCCCTGACAAACCC-----AATCTAGAGATACAAATGTAACGTCAAGTGG	29136
OY	2668	AGGGTTAGTGCAGATCTCAGAGATTATCAATGAGAGCCGTGTCTTTATACCAAGCTG	2727
Db	29137	A-GACTAGTGGAGATCTTAGCATCATAT-CAGGACGTAATTCACATATATCCAGTTG	29194
OY	2728	TACCTAGCCCTTATCTGTGTGTTCCCAATACCAGAGAGAGAGTGGTTTAAATGCC	2787
Db	29195	TACCAACCCCTTATCCCTGCTCTCAATATACAGAGAGAGAGATGGTTCACGGTTC	29254
OY	2788	TGACCTTTMAGAGATGCTTCTTGTCATCCCTGTATACCTGTCAATCTTGTTTG	2847
Db	29255	TGACCTCAAGAGATGCTTCTTGTTATCCCTGTCACTGTATCCAGTTCTCTTG	29314
OY	2848	CCTTTGAAGATCTTCAAAACCCACATCTCACTCACTGAGCTRTTTTATCCCAAGGT	2907
Db	29315	CCTTTGAAGATCCCAACCAACCAATCCCACTTTCAGTGGAGCGTCTTGCCCAAGGT	29374
OY	2908	TCAGGGTATATCCCATCTATTTGGCAGGCACTTACCCCAAGCTTGAGACATATTCAT	2967
Db	29375	TTAGGGATACCCCTCATCGGTGTGGTGAAGCCCTAGCCCAAGATCTAGGCACTTCTCA	29434
OY	2968	ACCTGGAACACTGTGTGCTTCRGTAKGTGATGTTACTTTTGGCGCYATTGAGAA	3027
Db	29435	GTCAGGCACTGTGTCTTCAATATATGTGATATTACTTTGGCTACAGTTGAGAG	29494
OY	3028	CTTGTGTCCATCAAGCAACCCCAAGCCTCTTMAATTTCCCGYACCTGTGCTCAAG	3087
Db	29495	CCTGTGTCCAGCAGGCTACTGTAGATCTTGAATCTTGAAGTATCAAGGATACAAAG	29554
OY	3088	TTTGCA-AASBARAGCTCARCTGTGCTCACACAGGTTAAATACTTAGGCTTAATTA	3146
Db	29555	TGTCTAGTGCAGAGCCCAAGCTTTGCTTACAGAGTTTAAATCTTAGGCCCTAATCTTA	29614
OY	3147	TCAAAGGCACARAGCCCTCATGAGGAAYATACAGCTTATCTGCTTATCTCTAT	3206
Db	29615	GCCAAAGGACAGAGGCCCTCATAGAAAGCAATACAGCTTATCTGCTTATCTCTGAC	29674
OY	3207	CYCAAAACCCTAAGCAACTPAGGRTTCTTGCGCTAAYAGSYTTTCGCGAATAATGG	3266
Db	29675	CCTAGACATTTAAACAGTGTGTGGGGTCTCTTGAAATTA-CTGGCTTTGGCGACTATGG	29734
OY	3267	ATTCCCAAGGWTGGCRBAATAGCCAGGYCAATTAMATACASTAATTAAGAAATCAGAA	3326
Db	29735	A-TCCCCGGAATACAGCGAGATAGCCAGGCCACTGTAGCCTATACAGGACACCAAG	29793
OY	3327	AGCCAAATCCCATTTATPAGATGGAYAMCTGAAGTTPAAGTGGCTTTCAGGCCCTTA	3386
Db	29794	GCGAAATAC-----TGTGGTGGGAATCAGAGCGAGAAACAGCCTTCAAAACCTTAA	29845
OY	3387	AGAAAGCCCTTAAACCAAGYCCCAAGTGTAAAGYTTGCAACRGGGCAAGACTTSTTAA	3446
Db	29846	GAGGGCCCT--AGTACAGCTCCAGCTTTAAGCTTCCACAGTACAAAGCTTCTTTTA	29903
OY	3447	TATATCAGAAAAAAAACGPAAYAGCTCTTGAAGTCTTTACACAGTCCRAGAGAYGAG	3506
Db	29904	TACGTACAG-AGAGAGCCGGGATAGCTCTTGAGAGTCTTATCTAGACTGTGGGACAAC	29962
OY	3507	CTTGCAACCTTGTGCRVACTGASTAAGGAAYATGATGTGGCAAAAGGATGRCYTCA	3566
Db	29963	CCCAACCAAGGCGATACCTTAATGAAATTAATGATATACAAAGGCTGGGCTCA	30022
OY	3567	TTGTTTAAAGGAGTAGTGAGCAGTAGCAGTATATGATCTGAAAGCACTTAAATTAATCA	3626
Db	30023	CTGTTTAATGTATGTGACAGCAGTGGCGGTCTTAGTGTACAGACATCAAAATTAATACA	30082
OY	3627	GGGAGAGATCTTACTGTGTGCACTCTCATAGAGTGAATRGCTACTCACTCTTAAAG	3686
Db	30083	AGGAAGAGATCTCACTGTCTGAGACTCAATGATATGAAATGGCACTATCTAGGTGCAAGG	30142

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```







Db 1125 CAGGGCCCTCAGTGAAGAGACGATATCAGCTTACTGCTTATCTCTCAATCCCAAAACCT 1184  
 Qy 3218 AAAGCACTAAGRGRRRTTCCTTGCGRTAA YAGTCTTTCGCCAAWATGATTCGCCAGGT 3277  
 Db 1185 AAAGCACTAAGAGGGTTCCTTGCGRTAA CAGTCTTTCGCCAAWATGATTCGCCAGGT 1243  
 Qy 3278 WTGGCAATAATGACGAGYCATTAATATACATAATTAAGGAACTGAGAAAGCCATAATCC 3337  
 Db 1244 ACAGCCCAATATGACGAGCATTAATATACATAATTAAGGAACTGAGAAAGCCATAATCC 1303  
 Qy 3338 ATTTATATATATGAGYAMCTGAGTAAAGTGGCTTTTCCAGGCCCTTAAGAGGCTTA 3397  
 Db 1304 ATTTATATATATGAGYAMCTGAGTAAAGTGGCTTTTCCAGGCCCTTAAGAGGCTTA 1358  
 Qy 3398 AACCAAGYCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3457  
 Db 1359 AACCAAGYCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1417  
 Qy 3458 AAAAAGCAAGAAATGCTTGGAGTCTTGAACAGATCCAGTCCAGAGGAGAGTGAAGTGAAG 3517  
 Db 1418 AAAAAGCAAGAAATGCTTGGAGTCTTGAACAGATCCAGTCCAGAGGAGAGTGAAGTGAAG 1477  
 Qy 3518 TGGCTTATATATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3577  
 Db 1478 TGGCTTATATATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1537  
 Qy 3578 TAGTGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3637  
 Db 1538 TAGTGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1597  
 Qy 3638 TTAATGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3697  
 Db 1598 TTAATGTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1657  
 Qy 3698 TGTGAGCAACATGTTTACTTAAATATCAAGTCTTAAATATCAAGTCTTAAATATCAAGTCT 3757  
 Db 1658 TGTGAGCAACATGTTTACTTAAATATCAAGTCTTAAATATCAAGTCTTAAATATCAAGTCT 1717  
 Qy 3758 TGTGAGCAACATGTTTACTTAAATATCAAGTCTTAAATATCAAGTCTTAAATATCAAGTCT 3817  
 Db 1718 TGTGAGCAACATGTTTACTTAAATATCAAGTCTTAAATATCAAGTCTTAAATATCAAGTCT 1776  
 Qy 3818 AAATATATCTGCAACATTAATTTCTCAAACTATGCACTGAGGGAGCTTGTAGAG 3877  
 Db 1743 AAATATATCTGCAACATTAATTTCTCAAACTATGCACTGAGGGAGCTTGTAGAG 1796  
 Qy 3878 TTCTTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3934  
 Db 1777 TTCTTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1836  
 Qy 3935 AAAGCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3991  
 Db 1837 AAAGCACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1896  
 Qy 3992 AAATCCCT---CACTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4045  
 Db 1897 AAATCCCT---CACTCAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1956  
 Qy 4046 KGGGCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4105  
 Db 1957 KGGGCACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2016  
 Qy 4106 YACCTAGTCAATC-----ATGCCCATGMRGCAATATG 4138  
 Db 2017 YACCTAGTCAATC-----ATGCCCATGMRGCAATATG 2076  
 Qy 4139 AAGGAAAGGAGATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCT 4198  
 Db 2077 AAGGAAAGGAGATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCTTAATCT 2136  
 Qy 4199 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4258  
 Db 2137 AATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2196

Qy 4259 CANAAGGAAAGRAAGGAAATASAGGAAATGCCAA 4297  
 Db 2197 CAGAAAGGAAAGGAGAGAGACACACATTAAGTCTTA 2235  
 RESULT 14  
 US-09-374-766-57  
 Sequence 57, Application US/09374766  
 Patent No. 6579526  
 GENERAL INFORMATION:  
 APPLICANT: HERVE PERRON  
 APPLICANT: FREDERIC BESEME  
 APPLICANT: FREDERIC BESEME  
 APPLICANT: GLAUCIA PARANOS-BACCALA  
 APPLICANT: FLORENCE KOMURIAN-PRADEL  
 APPLICANT: COLETTE JOLIVET  
 APPLICANT: BERNARD MANDRAND  
 TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS  
 TITLE OF INVENTION: ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL  
 NUMBER OF SEQUENCES: 92  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: OLIVIER & BERRIDGE  
 STREET: 700 South Washington Street, Suite 300  
 CITY: Alexandria  
 STATE: Virginia  
 COUNTRY: U.S.A.  
 ZIP: 22314  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/374,766  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/691,563  
 FILING DATE: 02-AUG-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Berridge, William P.  
 REGISTRATION NUMBER: 30,024  
 REFERENCE/DOCKET NUMBER: WPB 38588  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-836-6400  
 TELEFAX: 703-836-2787  
 INFORMATION FOR SEQ ID NO: 57:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2391 base pairs  
 TYPE: nucleotide  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 US-09-374-766-57  
 Query Match 18.7%; Score 1415.8; DB 4; Length 2391;  
 Best Local Similarity 78.4%; Pred. No. 0;  
 Matches 1817; Conservative 109; Mismatches 267; Indels 126; Gaps 19;  
 Qy 2021 ATGATCAGAGCAGAGCAGTGAAGTGCCTGGGGCAAGCGCCATCCATGCAATCACCCTC 2080  
 Db 1 ATGATCAGAGCAGAGCAGTGAAGTGCCTGGGGCAAGCGCCATCCATGCAATCACCCTC 60  
 Qy 2081 ACAGAGCCCTGGGTATGCTTGAACATTTAGAGGCGCAGAGAGT---TGTCTCTGAGACACT 2137  
 Db 61 ACAGAGCCCTGGGTATGCTTGAACATTTAGAGGCGCAGAGAGT---TGTCTCTGAGACACT 120  
 Qy 2138 GGTGCGGTCTTCTTACTCTTCTGCTGCTGCGGCAACTGCTCCAGATCTGTCACCT 2197  
 Db 121 GGTGCGGTCTTCTTACTCTTCTGCTGCTGCGGCAACTGCTCCAGATCTGTCACCT 180  
 Qy 2198 ATTCTGAGGGGGTCTTAAGACGGGAGTCACTAATTAATCTTATCCAGGCACTAAGTTA 2257

No + prior



Sequence 53, Application US/08979847B  
Patent No. 6582703  
GENERAL INFORMATION:  
APPLICANT: PERON, HERVE  
BESEME, FREDERIC  
BEDIN, FREDERIC  
PARANHOS-BACCALA, GLAUCIA  
KOMURIAN-PRADEL, FLORENCE  
JOLIVET-REYNAUD, COLETTE  
MANDRAND, BERNARD  
GARSON, JEREMY  
TUXE, PHILIP  
TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS  
ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL  
THERAPEUTIC PURPOSES  
NUMBER OF SEQUENCES: 210  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OLIFF & BERRIDGE, PLC  
STREET: P.O. BOX 19928  
CITY: ALEXANDRIA  
STATE: VA  
COUNTRY: USA  
ZIP: 22320  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/979,847B  
FILING DATE: 26-No. 6582703-1997  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BERRIDGE, WILLIAM P.  
REGISTRATION NUMBER: 30,024  
REFERENCE/DOCKET NUMBER: WPB 39046A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-836-6400  
TELEFAX: 703-836-2787  
INFORMATION FOR SEQ ID NO: 53:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2391 base pairs  
TYPE: nucleotide  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 53:  
US-08-979-847B-53  
Query Match 18.7%; Score 1415.8; DB 4; Length 2391;  
Best Local Similarity 78.4%; Pred. No. 0;  
Matches 1817; Conservative 109; Mismatches 267; Indels 126; Gaps 19;  
QY 2021 ATGATCCAGCAGCAGCAGCAGTGGGCTGCTGGGGCAAGCCGCATCCATCCATCACCCTC 2080  
DB 1 ATGATCCAGCAGCAGCAGCAGGAGGCTGCGCCGGGGCAAGCCGCATCCATCACCCTC 60  
QY 2081 ACAGACCCCTGGGTATGCTTGAACATTGAGGGCCAGGAAGT---TCTCTCCGAGCACT 2137  
DB 61 ACAGACCCCTCAGGTATGCTTGAACATTGAGGGCTCAGAAAGGTATGCTCTCGAGCACT 120  
QY 2138 GGTGGCGTCTTCTTAGTCTTACTCTTGTGTCGGGCAACATGTCCTCAGATCTGTACT 2197  
DB 121 GGGGAGCCTTCTCAGTCTTACTTCTGTCTCTGAGCAATGTCCTCAGATCTGTACT 180  
QY 2198 ATTCTGAGGGGCTCCTTAAGACGGGCACTCAGATGATCTTCTTCCAGCCACTAAGTTA 2257  
DB 181 GTCC---GAGGGGCTCCTAGAGCAGCCAGCTCAGATGATAC-TTCTCCAGCCACTAAGTTG 236  
QY 2258 TGAACCTGGGAGCTTTATCTTTTTCACATGCTTTTCTTAATTAATGCTTGAAGCCCACTA 2317  
DB 237 TG-ACGGGGAACCTTAACTCTTCCCATGCTTTTCTTAATTAATGCTGAAGCCCACTC 295

QY 2318 CTTGTGAGSAGAGACATTTCTAGCAAAAAGCAGGGCCATTATACCTGAACATAGAG 2377  
DB 296 TCTGTGGGAGAGACATTTCTAGCAAAAAGCAGGGCCATTATACATGTGAATATAGAG 355  
QY 2378 AAGGAACACCCGTTGTGTGTCCTGCTTGAAGAAATTAATCTGTAAGTCTGGCA 2437  
DB 356 AAGGAACAATGTTTGTGTGTCCTGCTTGAAGAAATTAATCTGTAAGTCTGGCA 414  
QY 2438 ACAGAAAGCAATATGAGAGAGCCAAAGATGCGCGCTGCTCAAGTTAACTAAAGGA 2497  
DB 415 ACAGAAAGCAATATGAGAGAG-CAGAGATGCGCGCTGCTGCTCAAGTTAACTAAAGGA 473  
QY 2498 TTCCATCTCTCTTCCCTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAAGAGAT 2557  
DB 474 TTCCATCTCTCTTCCCTACCAAGGAGTACCCCTCAGACCCCAAGGCCCAAGAGAT 533  
QY 2558 CCAAAAAGATTGTAAGGACTTAAAGCCCAAGGCTTATGAATCAATGCTGCTG 2617  
DB 534 CCAAAAAGATTGTAAGGACTTAAAGCCCAAGGCTTATGAATCAATGCTGCTG 593  
QY 2618 CAGTAATTCGTAAGTGAAGGAGCAGAAACCAAGTGAAGTGAAGGAGTTAGT 2677  
DB 594 CAGACTCC-----AATTTAGAGTAAGAAACCAAGCAGTGAAG-GATTAGT 645  
QY 2678 CAGATCTCAGATTTATCAATGAGGCGCTGTCTTTATACCAAGCTGTACTAGCC 2737  
DB 646 CAGAACTCAGATTTATCAAT-GAGGCTGTGTCTCTATACCAAGCTGTACTAGCC 704  
QY 2738 TTATCTGCTGCTTCCCAATATCCAGAGAGAGAGTGTATACASTCTGCACTTMA 2797  
DB 705 TTATCTGCTGCTTCCCAATATCCAGAGAGAGAGTGTATACASTCTGCACTTMA 764  
QY 2798 GAGTGCCTTCTCTGATCCCTGATCATCTGATCTGATCTGATCTGATCTGATCTG 2857  
DB 765 GAGTGCCTTCTCTGATCCCTGATCATCTGATCTGATCTGATCTGATCTGATCTG 824  
QY 2858 TACTTGAACCCAGATCTCACTCACTGATCTGATCTGATCTGATCTGATCTGATCTG 2917  
DB 825 TACTTGAACCCAGATCTCACTCACTGATCTGATCTGATCTGATCTGATCTGATCTG 884  
QY 2918 YCCCCATCTATTTGGCAGAGGCTTACCCCAAGCTGATCTGATCTGATCTGATCTG 2977  
DB 885 CCCCCATCTATTTGGCAGAGGCTTACCCCAAGCTGATCTGATCTGATCTGATCTG 944  
QY 2978 TCTTGTCTCTGATGATGATGATTTTCTTGTGCGCCCTTCAAGAACTTGTGCA 3037  
DB 945 TCTTGTCTCTGATGATGATGATTTTCTTGTGCGCCCTTCAAGAACTTGTGCA 1004  
QY 3038 TCAAGCCACCCAGATCTTAAATTTCTCGATCTGATCTGATCTGATCTGATCTG 3097  
DB 1005 TCAAGCCACCCAGATCTTAAATTTCTCGATCTGATCTGATCTGATCTGATCTG 1064  
QY 3098 ABAAGCTCARTCTGCTACAGAGGTTAAATCTAGAGCTAATATTCGAAGGCA 3157  
DB 1065 ABAAGCTCARTCTGCTACAGAGGTTAAATCTAGAGGCTAATATTCGAAGGCA 1124  
QY 3158 CARGGCTCAGTGAAGAAATATATCAAGGCTTACTGCTTACTCATCTCACTCACT 3217  
DB 1125 CARGGCTCAGTGAAGAAATATATCAAGGCTTACTGCTTACTCATCTCACTCACT 1184  
QY 3218 AAGCAACTAAGGRTTCTTGGCTTAATATATATATATATATATATATATATATAT 3277  
DB 1185 AAGCAACTAAGGRTTCTTGGCTTAATATATATATATATATATATATATATATAT 1243  
QY 3278 ATGAGCAAT 3337  
DB 1244 ATGAGCAAT 1303  
QY 3338 ATTAT 3397  
DB 1304 ATTAT 1358  
QY 3398 AACCAAGCCAGATGTTAAGTTGCAACRGGGCAAGACTTTTATATATATATATATAT 3457

```
DB 1359 AACCCAGGCCCGAGTGTCCAGCTTGCCACAGGCAAGATTTTCTTATATGCGCACAG- 1417
QY 3458 AAAAAACAGBAAVAGCTCTRGAGTCCCTTACACAGTCCBAGGAAVAGCTTCAACCYR 3517
DB 1418 AAAAAACAGBAAVAGCTCTRGAGTCCCTTACAGGCTCTCAGGGAATGAGCTTCAACCCG 1477
QY 3518 TGGCRYACCTGASTAAGAAAATGATGTAGTGGCAAGGGTGRCYTCATTGTTTAYGGG 3577
DB 1478 TGGTATACCTGAGTAAAGAAATGATGTAGTGGCAAGGGTGGCTCATNGTTTAYGGG 1537
QY 3578 TTAGTGTGCGATGACAGTCTKTAGTATCTGAAGCAGTTAAATATATACAGGGAAGATC 3637
DB 1538 TAAATGAGGCAATGAGCAGTCTTAAATCTGAGCAGTTAAATATATACAGGGAAGATC 1597
QY 3638 TTAAGTGTGAGCATCTCATGAGTGAAYRGCATCTCACGCTAAAGSAGACCTGTGGC 3697
DB 1598 TTNGCTGTGAGCATCTCATGAGTGAAYRGCATCTCACGCTAAAGSAGACCTGTGGT 1657
QY 3698 TGTGACACAACVGTCTTAAATCTGAGGCTCTATCTTGAARGGCCAGTGTGCRAC 3757
DB 1658 TGTGACACAACVGTCTTAAATCTGAGGCTCTATCTTGAARGGCCAGTGTGCRAC 1717
QY 3758 TGTGCACTGTGTGCACTCTTAAACCCAGTCTTCCAGACAATGAAGAAAGATAR 3817
DB 1718 TGGGCACTGTGTGCACTCTTAAAC----- 1742
QY 3818 AAYATAAGTGTCAACARATATTTCTCAACCTATGCGACCTGAGGGACCTTGTAGAR 3877
DB 1743 -----CAATCTATGCTGCCAGAGGATCTTTNTAGAG 1776
QY 3878 TTCCYTTGACTGATCCYAGCTTCACT--TGTATACTGATGAGAGTCTTTGTAGAA 3934
DB 1777 GTCCCTTACCAACCTGACCTCACTATATATATCTGATGAGAGTGTGTGTAGAA 1836
QY 3935 AAAGGACTTGAAGAGGGGTATGC--AGTGTGCTAGTATATGGAATATGGAAGT 3991
DB 1837 AAGGATTTACAAAGGAGATATNCCATAGTGTGTATGATMAAGCATGTAAGT 1896
QY 3992 AATCCCT---CACTCCAGGAATAGTGTGCTGACGCTGACGAACATAATAGCCYT---CAYT 4045
DB 1897 AAGCTCTTCCCGCCCGAGGACCGCCCGCTTACGAGAACTAGTGGCACTGACCCG 1956
QY 4046 KGGGCACTAGAAATTAGAGAAAGGAAAGGAAATATATATACAGACTCTATATAGCT 4105
DB 1957 CGAGCTTAAACTTTGAAAAGGAGAGATTAATGTGTATACAGATAGCAATAGTCT 2016
QY 4106 YACCTAGTCNTCC-----ATGCCCATGMRGCAATAGS 4138
DB 2017 TATCTAATCCGAATGCCCATGTGTATCTAATCCGAATGCCCATGTGTCAATAGG 2076
QY 4139 ABAAGAAAGGAATCTTAATCTTGYAGBAAACCTATCMAACATCAGAAAGCATTAGG 4198
DB 2077 AAGAAAGGAGATCTTAACCTCTGGGGAAACCCCATTAATACCAAGTTAATCATG 2136
QY 4199 AATTTATTAYTGGCGWGTACAGAAACCTARAGAGGTGAGAGTCTTACACTGCGGGGTCA 4258
DB 2137 GAGTTATTGACACAGTGCAGAAACCTCAAGAGGTGGAAGTCTTACACTGCGCAAGCAT 2196
QY 4259 CAAAAAGGAAAGGAAAGGAAATASAGRGAAYTGCCAA 4297
DB 2197 CAGAAAAGGAAAGGAGAGAGACATTAAGTGGCTA 2235
```